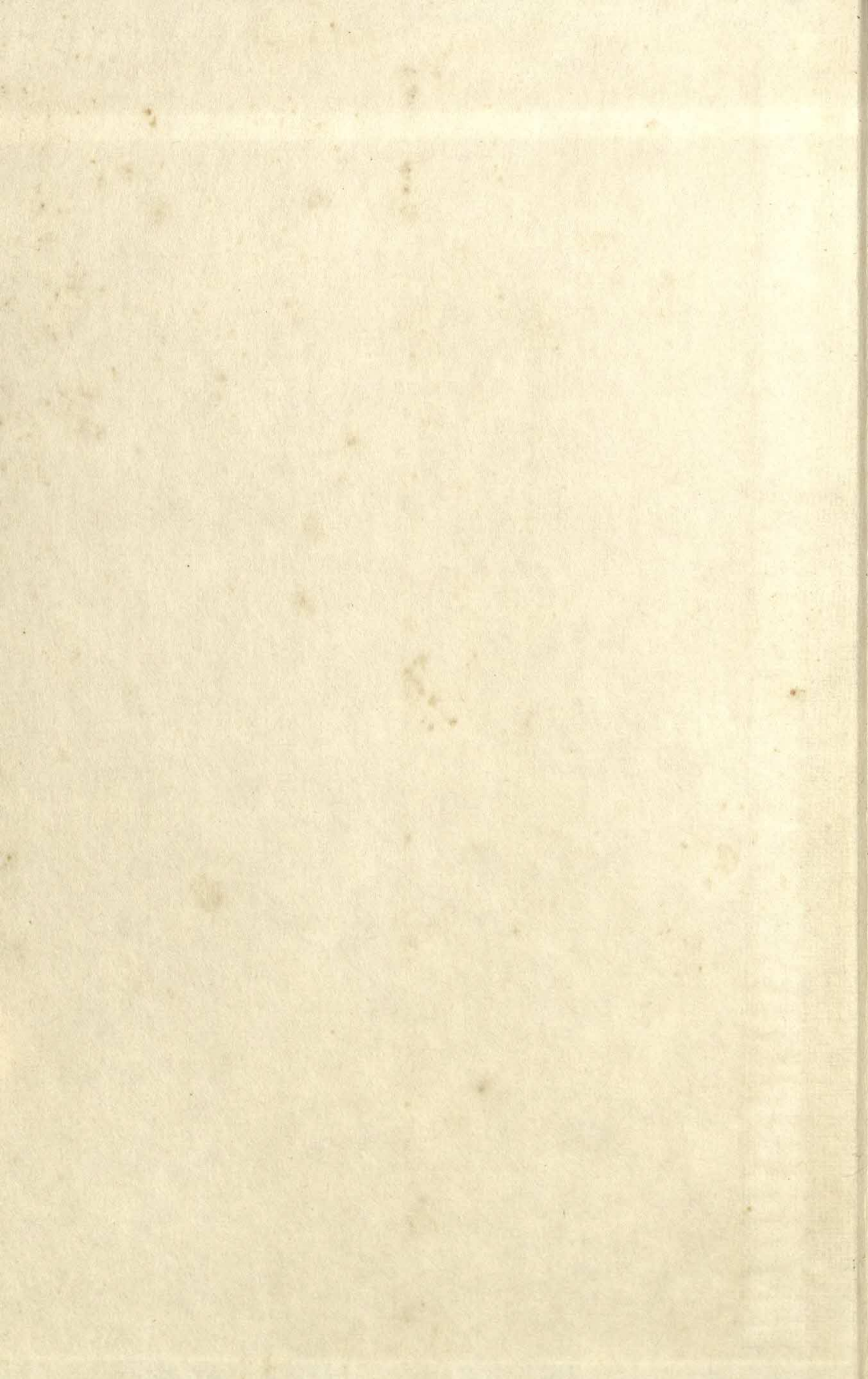


**GREAT
EXPERIMENTS
IN
PSYCHOLOGY**

HENRY E. GARRETT



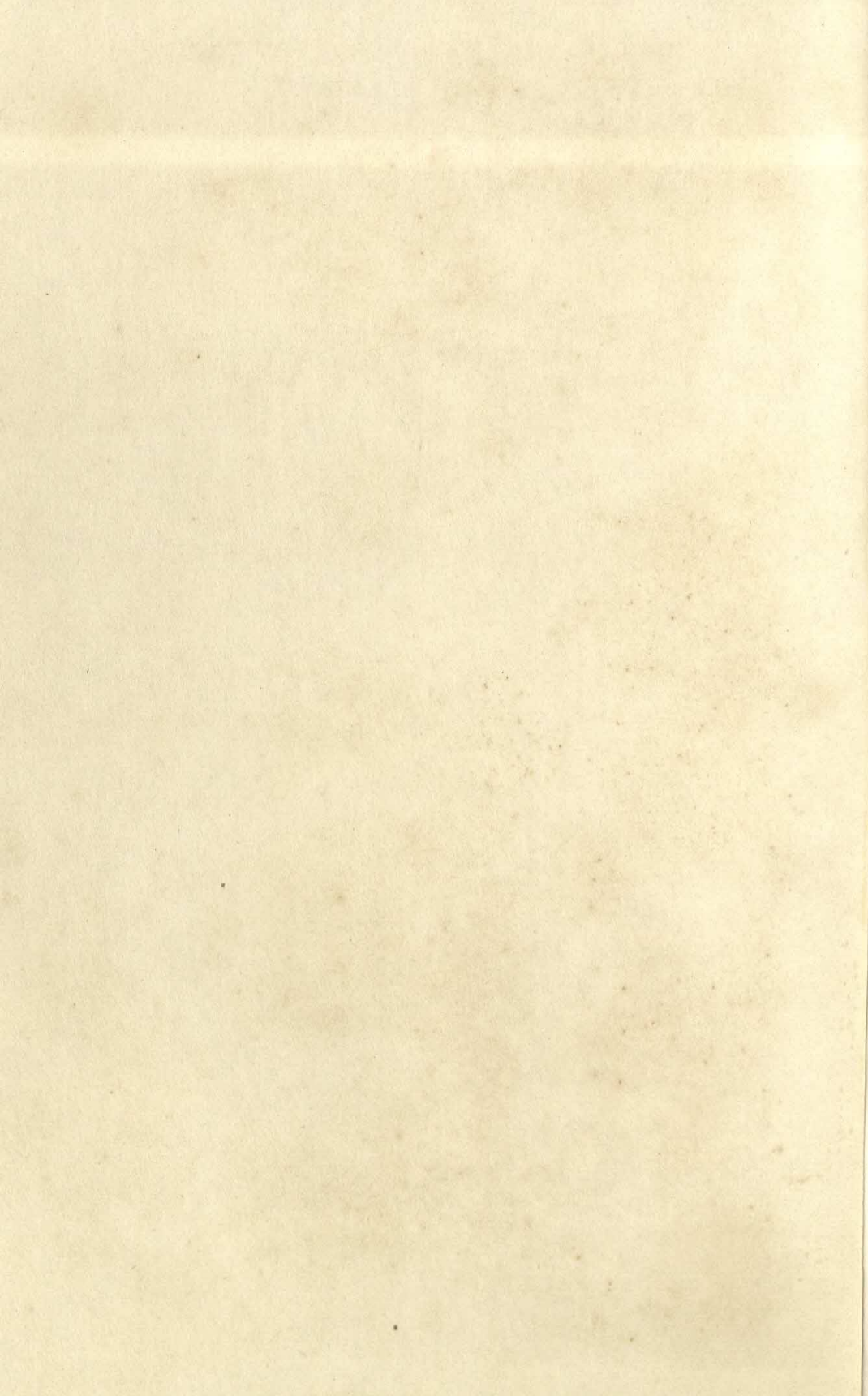
Psycho

SI-1697

~~D-456~~

~~4771~~





THE CENTURY PSYCHOLOGY SERIES

RICHARD M. ELLIOTT, *Editor*

KENNETH MACCORQUODALE, *Assistant Editor*



GREAT EXPERIMENTS IN PSYCHOLOGY

GREAT EXPERIMENTS IN PSYCHOLOGY

HENRY E. GARRETT

of

Columbia University



VISION

VISION PRESS LIMITED

Saxone House

74A Regent Street

London W1

LIBRARY
06.01.2009
13815

Published 1959

2nd Impression 1961

3rd Impression 1964

Bound in Great Britain

Printed in the USA

MCMLXIV

PREFACE TO THE THIRD EDITION

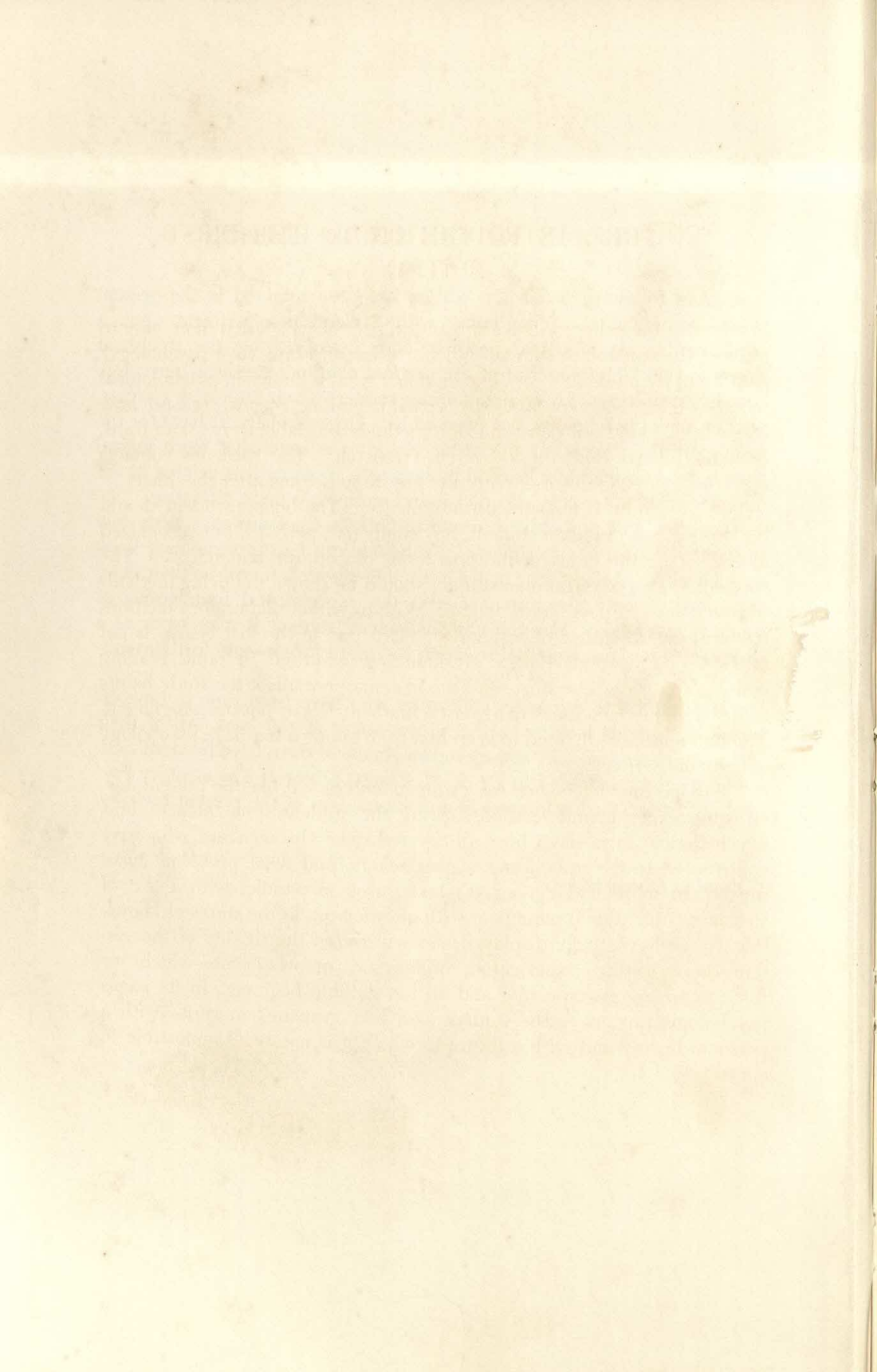
THE PLAN followed in the first edition has been retained in the present book—namely, that of beginning with a classical experiment upon a fundamental psychological problem and then bringing the treatment down to date. The material of the original chapters has been rewritten, revised, supplemented, and sometimes reinterpreted. One new chapter, that on social psychology, has been added. Titles of chapters have for the most part been kept but the order of topics is somewhat more logical than before. Instead of a general bibliography, reading lists printed at the end of each chapter provide quicker and more convenient reference.

The book will probably continue to find its chief use as a basic text in the first course in experimental psychology or as a supplementary book in general psychology. For many beginning students it will be their chief and perhaps only acquaintance with the experimental foundations of modern psychology. For the more advanced student it may serve as a quick review of fundamental research and as an incentive to further reading.

I am grateful to the many colleagues who have from time to time offered suggestions looking to the improvement of the book. All of their ideas have been carefully considered and many of them have been utilized. I am especially indebted to Dr. R. M. Elliott, Editor of The Century Psychology Series, for his interest in and stimulation of the present revision.

H. E. G.

Columbia University



EDITOR'S INTRODUCTION TO THE FIRST EDITION

Is THERE in the college curriculum any subject in which it is more easy to interest the student, and more difficult to instruct him, than psychology? At present the student is usually taught first what is known in psychological science (factual content, as in the typical beginning course), second, how it has been found out (experimental methods and laboratory), and third, how to discover something himself (research). The student's contacts with these three kinds of opportunity are stepwise, one after the other.

This procedure is unjustified traditionalism. The honors student should be treated as a graduate student, less competent perhaps, but recognized as employing the same mental operations of criticism and research. The sophomore or even freshman student should be shown the living methods of psychology, if possible in practice. With or without laboratory facilities to draw upon, the use of a single standard text in the first course is not to be recommended. Habits of skimming, acquired in rapid reading courses in history, literature, and the like, may not affect the study habits of the student when he is up against a mathematics or physics assignment. They do undoubtedly tend to keep him from taking a ten-page psychology assignment seriously.

There has long been need for a *supplementary* text like the present one, stressing experimental methods, giving the student some idea of how psychological facts have been discovered, who the men are who have contributed to the upbuilding of psychology, and what problems await immediate solution. Dr. Garrett's book, used in parallel with a text of standard content, will punctuate with question-marks the stores of knowledge so authoritatively displayed and will soften the rigidity of the system—behavioristic, experiential, organismic, or what not—which we force upon the unsuspecting and all too gullible beginner. In its pages psychology appears to the student as a live, growing enterprise with a personal history and with a future to which it is not at all impossible to contribute.

R. M. E.

THE HISTORY OF THE

UNITED STATES

OF AMERICA
FROM THE FIRST DISCOVERY OF THE COUNTRY
TO THE PRESENT TIME
IN TWO VOLUMES
BY
JOHN ADAMS
OF THE MASSACHUSETTS
IN 1795
LONDON: PRINTED BY J. JOHNSON, ST. PAULS CHURCH-YARD.
1795.

CONTENTS

	PAGE
PREFACE	v
 CHAPTER	
1 PAVLOV AND THE CONDITIONED REFLEX	1
1. The Conditioned Reflex	1
2. Pavlov's Method	3
3. How Conditioned Reflexes Are Established	4
4. The Factor of Time in Conditioned Reflexes	5
5. Sensory Discrimination and Generalization in Conditioned Reflex Experiments	7
6. Inhibition and Extinction	8
7. The Conditioned Reflex and Experimental Neurosis	10
8. The Conditioned Reflex and Sleep	12
9. The Conditioned Reflex in Animals and in Children	14
10. The Conditioned Reflex in Adults	16
11. The Conditioned Reflex as a Principle of Learning	17
Bibliography	18
✓ 2 FRANZ'S AND LASHLEY'S STUDIES OF THE ROLE OF THE BRAIN IN LEARNING	20
1. The Brain in Learned Activities	20
2. Early Studies of Brain Function and Learning	21
3. The Vicarious Rôle of Brain Structures: Restitution of Function	24
4. The Brain in Brightness Discrimination	27
5. Brain Function in Motor Habits	29
6. Equipotentiality and Mass Action of the Brain	32
7. Electroencephalography or the Study of "Brain Waves"	35
8. A Systematic Experiment in Psychosurgery	37
Bibliography	39
3 THORNDIKE'S LAWS OF LEARNING	40
1. Thorndike's Experiments in Animal Learning	40
2. Experiments with Fish and Chicks	41
3. Experiments with Cats, Dogs, and Monkeys	43
4. Do Animals Think or Reason in Solving Problems?	46
5. Do Animals Solve Problems by "Insight"?	50
6. Do Animals Use Imitation in Solving Problems?	52
7. Five Principles of Learning	53
8. Laws of Learning: The Law of Exercise	56
9. The Law of Effect	53
10. Present Status of the Law of Effect	59
Bibliography	61

4	KÖHLER'S EXPERIMENTS IN PERCEPTION AND LEARNING AND THEIR IMPORTANCE FOR GESTALT PSYCHOLOGY	63
	1. The Meaning of Gestalt Psychology	63
	2. Studies of Visual Perception	64
	3. Some Principles Governing the Formation of Perceptual Gestalts	66
	4. Learning from the Viewpoint of Gestalt Psychology	70
	5. Other Experiments on Learning and Motivation	73
	6. Studies in Gestalt Theory in America	76
	7. Present Status of Gestalt Psychology	78
	8. Insight and Educational Practice	80
	Bibliography	81
5	THORNDIKE'S AND WOODWORTH'S EXPERIMENTS ON THE TRANSFER OF TRAINING AND THEIR INFLUENCE UPON THE DOCTRINE OF FORMAL DISCIPLINE	83
	1. Formal Discipline and Faculty Psychology	83
	2. Evidence against Faculty Psychology	85
	3. First Experimental Attacks upon the Problem of Transfer of Training	87
	4. Transfer Experiments in Maze Learning	90
	5. Transfer Experiments in School Learning	91
	6. Theories of Transfer: Summary and Conclusions	95
	7. The Experimental Technique in Transfer Experiments	97
	Bibliography	99
6	EBBINGHAUS'S STUDIES IN MEMORY AND FORGETTING	101
	1. Ebbinghaus and the Invention of Nonsense Syllables	101
	2. Ebbinghaus's Experiments on Memory	102
	3. The Amount of Material to Be Memorized and the Time and Effort Required to Learn It	104
	4. The Amount or Degree of Learning and Retention	106
	5. The Time-Interval Between Learning and Recall	110
	6. Influence of the Method of Memorizing upon Retention	112
	7. The Retention of School Subjects	115
	8. The Effect of Repetition and Frequent Review upon Ability to Retain	116
	9. The Connections Formed in Learning	118
	10. Economy and Efficiency in Memorizing	120
	11. Summary of Ebbinghaus's Work	124
	Bibliography	125
7	WATSON'S STUDIES OF THE BEHAVIOR OF THE HUMAN INFANT	127
	1. Beginning of Experimental Study of the Child	127
	2. Motor and Sensory Development	128
	3. Mental and Social Development	131
	4. Handedness in Children	133
	5. Tests of Development in Young Children	135
	6. Emotional Development of the Young Child	139

CHAPTER

PAGE

7. How Children Acquire Fears and Other Emotional Attitudes	140
8. Methods of Breaking Undesirable Habits	143
Bibliography	146

8 CANNON AND EXPERIMENTAL STUDIES OF THE EMOTIONS 148

1. Psychological Theories of the Emotions	148
2. Problems and Methods in Studying Emotions	150
3. The Physical Basis of the Emotions	151
4. The Correlation of Physiological Changes and Emotional States	154
5. Judging Emotions from Facial Expressions	157
6. The Psychogalvanic Reflex	159
7. The Association Method in Emotional Difficulties	162
8. Attempts to Discover Guilt or Deception	166
9. Present Status of the Study of Emotions	168
Bibliography	169

9 EXPERIMENTAL APPROACHES TO PERSONALITY STUDY 171

1. The Meaning of Personality	171
2. Freud's Influence upon the Study of Personality	172
3. Factors in Personality	174
4. Impressionistic Methods of Studying Personality	175
5. Why Impressionistic Methods Have Vogue	181
6. Experimental Methods in the Study of Personality	183
7. Systematic Experimental Studies of Personality Traits	189
Bibliography	197

10 McDOUGALL AND SOCIAL PSYCHOLOGY 199

1. McDougall's Contribution to Social Psychology	199
2. Instincts in Social Psychology	200
3. Factors in Social Behavior	202
4. Methods in Social Psychology	207
5. Systematic Studies of Social Behavior	209
Bibliography	217

11 BINET AND THE INDIVIDUAL INTELLIGENCE TEST 219

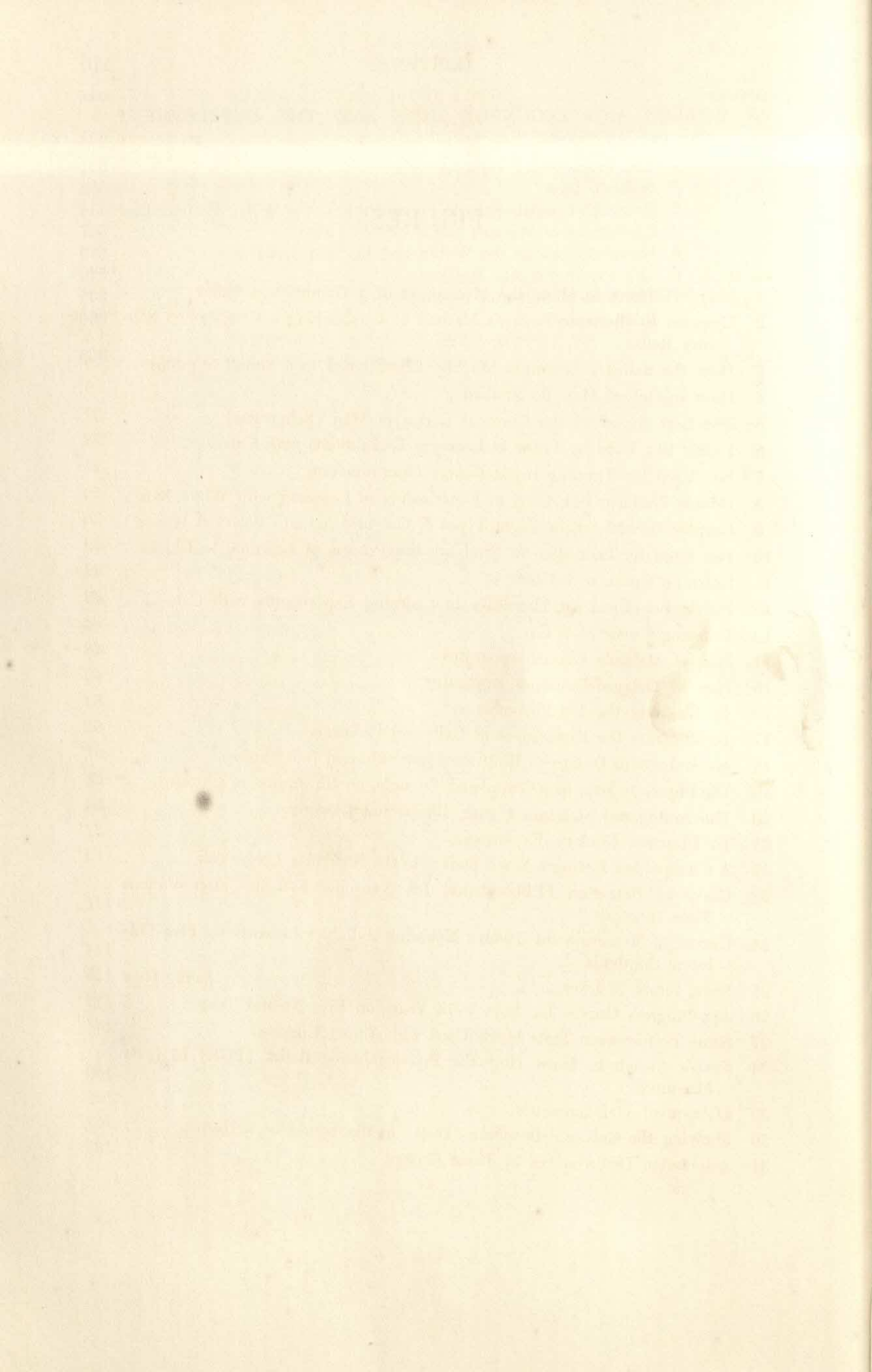
1. Beginning of the Movement for Measuring Intelligence	219
2. Development of the Binet-Simon Scale	220
3. Revisions of the Binet Scale in America	223
4. The Growth Curve of General Intelligence	225
5. Distribution of Intelligence	227
6. Intelligence Testing of Adults	229
7. Criticisms of the Individual Intelligence Test	232
8. The Inheritance of Intelligence	235
9. The Constancy of the IQ	239
10. IQ and School Success	240
Bibliography	241

12 THE ARMY TESTS: ALPHA, BETA, AND AGCT 244

1. The Army Tests of General Intelligence	244
2. Construction of a Group Intelligence Test	246

3. The Intelligence of the Soldier	249
4. Intelligence and Age	253
5. Intelligence and Occupational Level	254
6. Negro-White Differences on the Army Tests	257
7. National Differences in Intelligence	260
8. Intelligence and Schooling	261
9. Intelligence Tests in the Colleges	262
Bibliography	
13 GALTON AND THE MEASUREMENT OF INDIVIDUAL DIFFERENCES	264
1. The Beginning of Interest in Individual Differences	264
2. The Nature of Individual Differences	266
3. Correlation in the Study of Individual Differences	269
4. Nature and Nurture in the Study of Individual Differences	272
5. Statistical Studies of Inheritance	273
6. Further Studies of Nature and Nurture	275
7. The Study of Twins	277
8. Sex Differences	279
9. Race Differences	283
10. Natio-Racial Differences	287
11. Summary on Individual Differences	289
Bibliography	289
14 VISUAL AND AUDITORY PERCEPTION: THE WORK OF HELM-HOLTZ AND OTHERS	292
1. The Auditory Apparatus	292
2. How Sounds Are Located in the Environment	294
3. Theories of How We Hear	296
4. Visual Perception of Distance and Depth	297
5. How We See Objects in Three Dimensions	298
6. Physiological Factors upon Which Perceptions of Distance and Depth Depend	302
7. Psychological Factors in Distance and Depth Perception	304
8. Experience and Learning in Perception	306
9. The Nature of Eye Movements	307
10. Eye Movements and Other Factors in Reading	309
Bibliography	310
15 CATTELL'S STUDIES OF REACTION TIME	312
1. Reaction Time and Psychological Research	312
2. The History of the Reaction Time Experiment	313
3. Early Studies of Reaction Time	314
4. Apparatus Used in Measuring Reaction Time	316
5. Factors Influencing Simple Reaction Time	318
6. The Effect of Attitude upon Reaction Time	321
7. Associative Reactions	323
8. The Discrimination Time Method	324
9. Some Practical Applications of Reaction Time Studies	326
Bibliography	329

CHAPTER	PAGE
16 WEBER'S AND FECHNER'S LAWS AND THE DEVELOPMENT OF PSYCHOPHYSICS	331
1. The Problems of Psychophysics	331
2. Weber's Law	332
3. Fechner's Contributions to Psychophysics: the Weber-Fechner Law	334
4. Limitations of Weber's and Fechner's Laws	337
5. Interpretations of the Weber and Fechner Laws	339
6. The Psychophysical Methods	340
7. The Present Status of Psychophysics	349
Bibliography	350
INDEX	353



FIGURES

FIGURE	PAGE
1 Simple Scheme to Show the Mechanism of a Conditioned Reflex . . .	2
2 Diagram to Illustrate Pavlov's Method of Establishing a Conditioned Salivary Reflex	5
3 How the Salivary Secretion May Be Conditioned to a Sound Stimulus . .	6
4 How Inhibition May Be Studied	9
5 The Left Aspect of the Cerebral Cortex in Man (Schematic)	22
6 Puzzle Box Used by Franz in Learning Experiments with Cats	23
7 Box Used for Training in Brightness Discrimination	27
8 Double-Platform Box Used in Experiments in Learning with White Rats .	30
9 Graphic Records of the Eight Types of Cerebral Action—Potential Waves .	36
10 Pen Used by Thorndike in Studying the Process of Learning in Chicks .	42
11 Learning Curve of a Chick	42
12 Puzzle Box Used by Thorndike in Learning Experiments with Cats . .	43
13 Learning Curve of a Cat	44
14 Plan of Multiple Choice Apparatus	48
15 Plan of Delayed-Response Apparatus	49
16 To Illustrate the Phi-Phenomenon	64
17 To Illustrate the Experiment of Gelb and Goldstein	66
18 An Ambiguous Design to Illustrate Figure-Ground Relationship	67
19 The Figure Is Seen as a Completed Triangle, an Illustration of Closure .	68
20 The Ambiguous Staircase Figure, Illustrating Closure	69
21 To Illustrate Köhler's Experiment	71
22 A Chimpanzee Fitting a Small Stick into the Bore of a Larger One . .	73
23 Curve of Retention (Ebbinghaus) for Nonsense Syllables after Various Time Intervals	111
24 Curves of Retention for Twelve Nonsense Syllables Learned by Five Different Methods	114
25 Some Infant Reflexes	<i>facing page</i> 129
26 Age-Progress Curves for Boys 9-14 Years on Five Mental Tests . . .	132
27 Some Performance Tests Much Used with Young Children	137
28 Simple Sketch to Show How the Psychogalvanic Reflex (PGR) May Be Measured	160
29 D'Arsonval Galvanometer	160
30 Showing the Galvanic Response (PGR) to the Sound of a Horn . . .	161
31 Association Disturbances in Three Groups	165

FIGURE	PAGE
32 The "Powers and Organs of the Mind"	177
33 Human Somatotypes	179
34 Graphic Rating Scale	184-5
35 Dickens Motorists' Behavior at Stop Sign and Cross Traffic Corners	206
36 Class Structure, First Grade	211
37 Curves of Attitudes of School Children toward the Chinese Before and After Seeing a Pro-Chinese Film	212
38 Growth of Unwillingness to Continue the War	215
39 Mental Growth Curves for Superior, Normal, and Dull Children from Birth to Maturity	226
40 Distribution of Intelligence (IQ's) in the General Population According to the 1937 Revision of the Stanford-Binet Scale	227
41 Changes in Mental Ability with Age	231
42 Profile of a Subject on the Full Scale of the Wechsler-Bellevue Intelligence Scale	232
43 Distribution of Scores Made by 9,339,289 Men on AGCT	252
44 Decline in Intelligence Test Scores After Maturity	254
45 AGCT Score in Relation to Civilian Occupation	255
46 Predicting Success of Officer Candidates from Their Grades in the AGCT	258
47 Distribution of Brain Weights of 416 Male Adults	267
48 Distribution of Scores on a Vocabulary Test Made by 360 Eleven-year-old Boys	268
49 A Correlation Diagram Showing the Relationship of Height in Parents to Height in Offspring	270
50 Overlapping of Frequency Distributions	280
51 Sound Cage	295
52 Direction Finder	296
53 How the Image of an Object AB is Focused on the Retina	298
54 Stereoscopic Slide Drawn to Give Perception in Three Dimensions	299
55 The Principle of the Refracting Stereoscope	300
56 Brewster's Hand Stereoscope	301
57 The Doubling of "Near" Objects in Far Fixation	303
58 The Doubling of "Far" Objects in Near Fixation	304
59 The Müller-Lyer Illusion	306
60 Photographic Record of the Eye's Movement in Tracing the Outline of a Circle	309
61 The Hipp Chronoscope	317
62 Test Unit of the SAM Discrimination Reaction Time Test (CP611D2)	328
63 Graphic Representation of the Relation Between Increases in the Intensity of the Stimulus and Increases in the Intensity of the Sensation	335
64 Psychometric Functions Obtained with Two Judgment Categories	346
65 Some Apparatus Used in Psychophysical Experiments	347

ILLUSTRATIONS

	<i>facing page</i>
IVAN P. PAVLOV	2
KARL S. LASHLEY	22
EDWARD L. THORNDIKE	42
WOLFGANG KÖHLER	64
ROBERT S. WOODWORTH	84
HERMANN EBBINGHAUS	102
JOHN B. WATSON	128
WALTER B. CANNON	150
SIGMUND FREUD	172
WILLIAM McDUGALL	200
ALFRED BINET	220
LEWIS M. TERMAN	221
FRANCIS GALTON	266
HERMAN LUDWIG F. VON HELMHOLTZ	294
JAMES McKEEN CATTELL	314
GUSTAV THEODOR FECHNER	332

GREAT EXPERIMENTS IN PSYCHOLOGY

1

PAVLOV AND THE CONDITIONED REFLEX

1. THE CONDITIONED REFLEX

THE RESEARCHES on the conditioned reflex of the great Russian scientist, Ivan Pavlov, date from about the year 1904, when Pavlov was awarded the Nobel prize in medicine. Pavlov's early studies were published only in Russian, and for some time were little known to scientists of other countries. When finally they became available they were seized upon with avidity in America, especially by those psychologists dissatisfied with the introspective psychology then widely current. Here at last, it seemed to some, was the foundation upon which could be constructed a straightforward and naturalistic account of human behavior. So it happened that Pavlov's methods became the mainstay of behaviorism, the modern revolt against introspective psychology, and are today an important weapon of the strictly objective psychologist.

Pavlov's researches have broad and important implications for several areas of psychology. They have thrown light upon the mechanics of learning and habit-formation; have stimulated much research upon the problems of motivation; and have led to an ingenious theory of the age-old riddle of sleep. In addition, Pavlov and his associates have succeeded in producing abnormal mental states in animals akin to those widespread ailments called neuroses, which are thought by many physicians and sociologists to be, in part at least, a by-product of our present-day intense and complex civilization. Until Pavlov's death in 1936, at the age of eighty-seven, this remarkable man, the son of an obscure peasant priest, enthusiastically carried on his work aided by a devoted group of students and followers. Experimental reports from his laboratory have been read by scientific men the world over.

Pavlov's researches began with a well-known and frequently observed fact. We have all seen saliva drip from a dog's mouth as he waits in eager anticipation for his food. Probably most of us, too, have felt our mouths "water" at the sight or odor of appetizing food. Originally, i.e., in young animals or very young children, saliva does not flow until food is actually

into the mouth: it is an automatic and reflexive glandular reaction, the adequate stimulus for which is food-in-the-mouth. In time, however, as we know, the sight or the odor of food alone is able to initiate the saliva flow. Such a response, brought about by something other than its original, and biologically adequate, stimulus is a "conditioned reflex." The associated stimulus—the sight or odor of food—is called the conditioned stimulus, and the response to the conditioned or substitute stimulus is the conditioned reflex. The mechanism of the conditioned reflex is shown graphically in Figure 1.

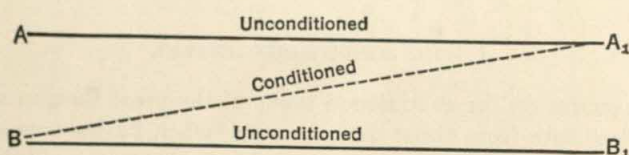


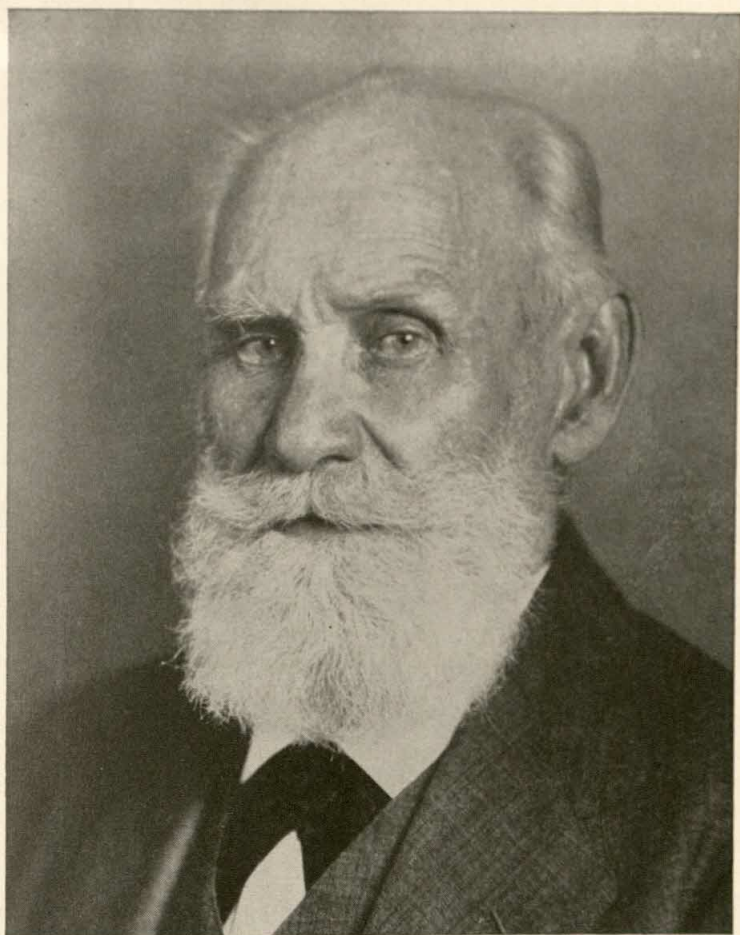
FIG. 1

SIMPLE SCHEME TO SHOW THE MECHANISM OF A CONDITIONED REFLEX

Call *A* the biologically adequate or unconditioned stimulus for the response A_1 ; and let *B* represent the biologically adequate stimulus for response B_1 . If *B* becomes an effective stimulus for the reflex A_1 , *B* is the substitute stimulus and $B \rightarrow A_1$ is the conditioned reflex. To illustrate, think of *A* as a clap of thunder which leads to A_1 , a startle or fear response; while *B* (lightning) leads to the mild response B_1 (blinking, say). Now if *B* accompanies *A*, as it usually does, *B* may also induce A_1 so that the fear response now follows the flash of lightning. Many of our reflexive activities are capable of being conditioned in this way. The salivary reflex is particularly well adapted for experiment, because responses to other than the original stimulus (i.e., food) can be accurately ascertained and their magnitude determined by noting the amount of saliva secreted.

The conditioned reflex has been widely employed in modern psychology to explain how we acquire new responses. But it would be a mistake to conclude that the term conditioned reflex or conditioned response (CR)¹ represents a new explanatory principle in psychology. In reality the idea of "conditioning" is distinctly old, and is clearly implicit in the time-honored laws of association by similarity, contrast, and contiguity in space and time. Essentially what the laws of association attempted to do was to explain how one idea or thought grows out of, or is connected with, or substitutes for, another. As early as 1690, John Locke in his *Essay*

¹CR usually stands for *conditioned response*. This term is more general than *conditioned reflex*, since the term *reflex* is ordinarily restricted to involuntary and unlearned responses (muscular or glandular) to sensory stimuli.



Bettmann Archive

IVAN P. PAVLOV

1849-1936

Concerning Human Understanding gave illustrations to show how individual peculiarities, likes, and dislikes can be explained on the principle of association. Locke describes two instances as follows: "A grown person surfeiting with honey no sooner hears the name of it, but his fancy immediately carries sickness and qualms to his stomach . . ." "A friend of mine knew one perfectly cured of madness by a harsh and very offensive operation. The gentleman who was thus recovered, with great sense of gratitude and acknowledgment owned the cure all his life after . . . but he could never bear the sight of the operator." Much later (in 1890) William James (Vol. 1, p. 566) stated the law of contiguity, to which the other laws of association may be reduced, in the following passage: "When two elementary brain processes have been active together or in immediate succession, one of them on recurring, tends to propagate its excitement into the other." This is substantially the same *kind* of explanation of association as that which Pavlov gave of conditioning.

If the idea of the conditioned response is not new, how, one may ask, does the work of Pavlov differ from that of the older association psychologists? The answer essentially is that Pavlov dealt not with "ideas," nor with complex verbal learning, but with glandular responses which can be measured and compared. The great merit of Pavlov's method is that responses were always measured under controlled conditions. Without measurement there can be, ordinarily, no definiteness, no accuracy, and no objective fact. Without definiteness, accuracy, and objective fact there can be no science.

In the following sections we shall first describe Pavlov's work with the conditioned reflex. Some of the broader aspects of conditioning and the rôle of the CR in learning will then be discussed.

2. PAVLOV'S METHOD

In his studies of the conditioned reflex, Pavlov worked almost entirely with dogs. How fast a dog can learn, and how rapidly he unlearns was studied by measuring the saliva flow under rigidly controlled conditions. As was said above, Pavlov chose to work with the salivary reflex because the magnitude—or the degree—of a response, and not simply its occurrence or non-occurrence can be determined from the *amount* of saliva secreted. Besides, the salivary glands form a simple organ and not a composite one consisting of several muscles; there are no muscular tensions present to interfere with or complicate the experimental control; and the secretion can be measured with great precision in units as small as one tenth of a drop.

Pavlov's method of collecting the saliva was relatively simple. By means

of a small incision, a fistula or opening was made in the dog's cheek through which a glass tube was inserted into the opening of one of the salivary glands, i.e., either the parotid, or the submaxillary. The saliva which dripped from the tube was collected and measured in finely graduated containers. The salivary reflex is delicate and readily interfered with, and the greatest precautions were taken in Pavlov's laboratory against any disturbances. All experiments were conducted in especially built windowless sound-proof rooms (there were eight in all), the walls of which were of turf, two feet thick. Both the food and the substitute stimuli were presented automatically, the food by an ingenious pneumatic device. Meanwhile the experimenter watched the dog from another compartment through a periscope in the wall, so that by no chance could the dog respond to him rather than to the stimuli. Each dog was put through a long training period, which sometimes lasted many weeks, until he was thoroughly accustomed to the surroundings and to the man who was to work with him. He was taught to stand upon a small table in the experimental room on which he was secured by a collar and other restraining but not uncomfortable bonds. It is reported that the dogs looked forward with every appearance of eagerness to the experiment, jumped upon the table without command and placed themselves in the correct position for the experiment. A diagram of the laboratory setup is shown in Figure 2.

3. HOW CONDITIONED REFLEXES ARE ESTABLISHED

The simplest technique in conditioned reflex experiments is to apply over and over again the conditioned stimulus together with the unconditioned or natural stimulus. In Pavlov's experiments, olfactory, auditory, visual and tactual stimuli were all used as substitute stimuli, food in every case being the unconditioned or adequate stimulus. With the hungry dog standing more or less quietly on the table, a small dish of food was presented, and at the same time a tone, say, or a buzzer sounded. This was repeated time after time on successive days, the number of joint stimulations averaging from eight to ten per experiment. At first the saliva flowed only to food *plus* sound, but finally, after repeated joint stimulations, usually from ten to twenty, the saliva began to flow at the sound alone. The sound is now a substitute stimulus for food, and the dog's salivary response is said to be conditioned. Figure 3 shows schematically how this change in stimulus-response relationship comes about.

Other stimuli than auditory have been substituted for food and made to produce the saliva flow—olfactory stimuli, such as the odor of camphor; visual stimuli, such as letters and geometrical forms; tactual stimuli,

touches or light scratches on the skin. In all of these experiments the dog may be said to have formed a connection between the "new" stimulus (e.g., odor) and the response (salivary flow), or to have "learned" that the conditioned stimulus "stands for" the food. Through associations like these, the range of stimuli which may lead to a given simple reflexive act becomes greatly increased (p. 53).

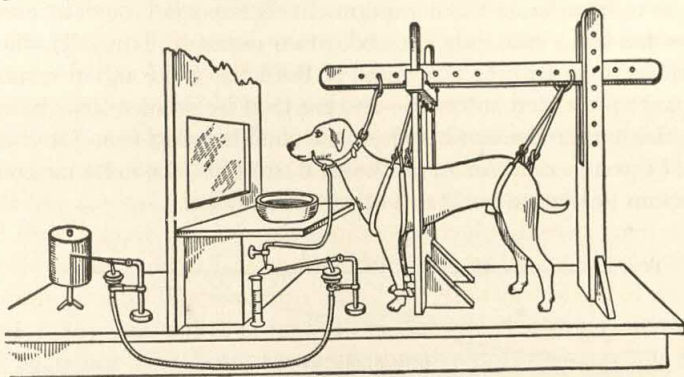


FIG. 2

DIAGRAM TO ILLUSTRATE PAVLOV'S METHOD OF ESTABLISHING A CONDITIONED SALIVARY REFLEX

The unconditioned stimulus (food) was presented automatically in the small dish through the window. At the same time, or prior to this, the conditioned stimulus, e.g., the ringing of a bell, was presented. The saliva which flowed from the dog's mouth was collected in a graduated glass receptacle. As the saliva flowed into the receptacle, it struck a small disk which depressed the lever just in front of the animal. This downward movement was transmitted to the lever behind the screen, and an automatic tracing was thus secured upon a smoked drum or kymograph. The kymographic record tells the experimenter how many drops of saliva have been secreted and how regular the flow has been. (From Yerkes and Morgulis, 1909.)

4. THE FACTOR OF TIME IN CONDITIONED REFLEXES

Time is an important factor in conditioned reflex experiments. When the conditioned stimulus is given *simultaneously* with the natural stimulus, the conditioned reflex can always be obtained. Conditioned reflexes may be obtained, also, when the conditioned stimulus is applied *before*² the natural stimulus—times as short as one second and as long as five minutes have given positive results. But if the conditioned stimulus comes *after* the natural stimulus for even so short a period as one second there is ordinarily no conditioned reflex. Krestovnikov, one of Pavlov's associates, carried out 1,000 trials in which the conditioned stimulus—a scratch on the skin—was applied from 1 to 3 seconds *after* the unconditioned stimulus (food) without getting positive results, i.e., saliva flow.

² Near-simultaneity (1 to 3 seconds *before*) is often better than strict simultaneity.

If a conditioned stimulus (e.g., tone or touch) is *continued* until the unconditioned stimulus is supplied the dog may, after considerable training, reach the point where its *latent time*³ of response approximates more and more closely the time between the application of the conditioned stimulus and the natural stimulus. Thus, if a note is sounded steadily on a horn, and two minutes later food is given, the saliva will not begin to flow, once the conditioned reflex has been established, until the horn has been sounding for about two minutes. This is called a "delayed reflex." Apparently, the sound of the horn is the *signal* which starts off the dog's nervous machinery—and the time between adequate and substitute stimulus determines how long the machine shall run. The longer the interval between conditioned and natural stimulus, the more numerous the applications needed to build in a conditioned reflex.

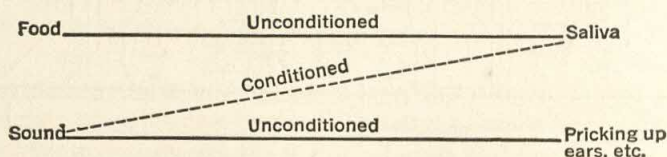


FIG. 3

HOW THE SALIVARY SECRETION MAY BE CONDITIONED TO A SOUND STIMULUS

When the conditioned stimulus is applied before the natural stimulus (food) it is not necessary that it be continued until food is presented. The conditioned stimulus may first be applied, a pause allowed, and the food then presented. The pause between the conditioned and the natural stimulus may be of considerable length; even three minutes is not unusual. The latent time in these experiments approximated the interval between application of the substitute stimulus and the food. For example, a note was sounded once, and two minutes later food given, and this process was repeated many times. Eventually, when the animal was conditioned, saliva began to flow about two minutes *after* the sound; that is, close to the time at which the food ordinarily appeared. This type of delayed response is called a "trace reflex." Pavlov believed that the intervening period was not simply one of inactivity since any extra stimulus (of optimal strength) which by chance or otherwise was applied during the waiting period caused an immediate secretion of saliva.⁴ The restraining mechanism at work during this wait is apparently like a delicate balance which can be upset by any strong supervening stimulus.

Another phenomenon closely akin to the trace reflex is called simply

³ The time between the application of the stimulus and the appearance of the conditioned reflex is the latency or latent time of the response.

⁴ See also p. 9.

the "time reflex." If a dog is fed at stated intervals, say every ten minutes, he soon becomes conditioned to this time-interval so that saliva will flow every ten minutes in anticipation, as it were, of food. It would seem in such cases that a rhythm or periodicity has been established in the organism which "sets off" the response at the proper time, much as a repeater alarm clock rings every two minutes or so. Examples of such rhythms or time reflexes in everyday life at once come to mind. Some people awaken at the same time every morning, oftentimes just before the alarm clock rings, while many workers through vague but unmistakable inner cues, know when their regular lunch time has arrived without consulting their watches. The fact of periodicity in efficiency during the day—physiological CR's possibly—seems to be well established.

5. SENSORY DISCRIMINATION AND GENERALIZATION IN CONDITIONED REFLEX EXPERIMENTS

Closely bound up with the question of the time relations between substitute and natural stimulus is that of the degree of differentiation possible between various coacting stimuli. This is clearly an important problem, for obviously there must be discrimination and choice among stimuli. Otherwise, the dog's salivary secretion would be free and uncontrolled, since dogs are continuously being bombarded by a multiplicity of sounds, smells, sights, and touches while eating. The typical sequence in conditioning experiments is generalization followed by differentiation. At first the dog may respond indiscriminately with saliva flow to several equivalent stimuli; later, through reinforcement of the one and non-reinforcement of the others, a choice is accomplished. How is this selection made and how fine is the dog's discriminative ability? Pavlov and his workers attacked the problem in several experiments, one of which is described here. First, two spots, *A* and *B*, on the dog's flank were selected. The *A* spot was lightly scratched, and at the same time food was given. Then *B* was lightly scratched, but no food was given. After a number of repetitions, sometimes spaced out over a period of days and even of weeks, *A* with food, *B* without food, the point was reached where touching *A* always gave the salivary reflex, and touching *B* never gave it. The negative spot *B* was then moved closer and closer to the positive spot *A*, in order to test the fineness of discrimination, until the two spots were separated by only a few millimeters. Still *A* gave the reflex, while *B* invariably failed to give it. Inevitably, however, this very fine discrimination broke down when a certain point was reached. Whenever the negative stimulus was almost identical with, or close to, the positive stimulus, both produced the conditioned reflex, the degree of response (amount of saliva) in the case of the negative stimulus

depending upon its nearness or similarity to the positive stimulus. This phenomenon of stimulus spread or generalization was attributed by Pavlov to "irradiation" in the cerebral cortex of the dog. The extent of stimulus generalization was found to be somewhat reduced by frequently reinforcing the positive spot with food and by not reinforcing the negative spot.

Clear-cut differentiation of stimuli has been secured by Pavlov's workers with auditory and visual as well as with tactual stimuli. These experiments throw light upon the extreme sensitivity of the dog's nervous system. Beliakov, for instance, trained his dogs to discriminate between tones of 800 and of 812 and of 825 vibrations per minute by reinforcing the positive stimulus and not reinforcing the others. He found that one must begin with the least similar (825 vibs.) and proceed to the more similar (812 vibs.) tone in order to get discrimination from 800, as the reverse procedure led only to generalization. In other experiments, it was found that the dog could distinguish between metronome beats at the rate of 96 and 100 per minute; and in the case of visual stimuli between a circle and an almost circular ellipse (i.e., axes in the relation 7:8).

6. INHIBITION AND EXTINCTION

The ability of the brain and nervous system of the dog to distinguish between two closely similar stimuli, sounds, touches or visual objects, was attributed by Pavlov to an actively restraining neural mechanism called *inhibition*. Inhibition is roughly analogous to the braking of an automobile which is gathering too much speed on a hill. This concept is well illustrated by an experiment in which four spots, A_1 , A_2 , A_3 , and A_4 , were selected on a dog's flank (Fig. 4). When each of these spots was touched, food was given, and the process was continued until all of the spots called out the salivary reflex. These may be called *positive* spots and the response is a common or generalized one. Another spot, B , in the same region, was next selected and touched again and again, no food being given on these successive stimulations. The procedure of presenting conditioned stimuli regularly without reinforcement (food) was called by Pavlov "experimental extinction" (see p. 10). At first, touches on the B spot led to a saliva flow due to stimulus generalization; but finally, as no food was ever given, the conditioned reflex ceased, and the spot became *negative*. Now, when the negative B was touched and afterwards one of the positive A spots, an interesting result occurred. Spot A_4 gave no response, A_3 gave about one half its usual quota of saliva, while A_2 and A_1 gave full effect. This highly interesting result was explained by Pavlov as due to the spread of the inhibiting or restraining action set up in the dog's nervous system. Clearly, the inhibition was related to the distance of the A spot from the negative

B spot. Such inhibition of a positive by a negative stimulus is a general fact often encountered in conditioning experiments—it occurs with auditory, visual, and olfactory, as well as with tactual stimuli.

Ordinarily, as we have seen, the negative *B* spot did not give a saliva flow. But if when *B* was touched, a bell or note was sounded at the same time, the conditioned reflex at once appeared—saliva began to drop as though *B* were a positive *A* spot. This lifting of an inhibition or *disinhibi-*

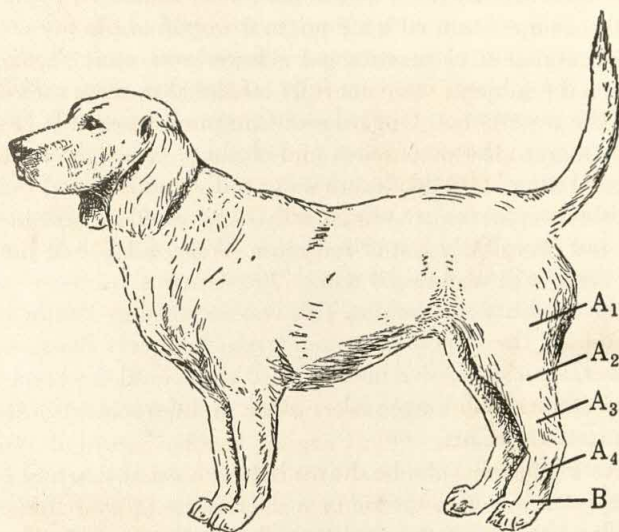


FIG. 4

HOW INHIBITION MAY BE STUDIED

Spots *A*₁, *A*₂, *A*₃, and *A*₄ are positive spots; *B* is a negative spot.

tion is also fairly common in conditioned reflex experiments; it is demonstrated most clearly, perhaps, in the trace reflex (p. 6). It will be remembered that in the trace reflex the salivary response did not appear until a given time *after* the substitute stimulus, the exact time depending upon the period originally intervening between the substitute stimulus and the food. During the waiting period the reflex is apparently held in check or inhibited by the dog's nervous system until the proper time arrives for releasing the saliva. But any strong extraneous stimulus will lift the brake, causing the salivary reflex to appear. For example, in an experiment conducted by Anrep, one of Pavlov's students, an irritating odor caused twenty-eight drops of saliva to flow during a trace reflex; while on another occasion the buzzing of a fly had the same disorganizing effect. Perhaps the reader will now understand better the need for soundproof, windowless rooms, and for the extreme precautions taken in conditioned reflex work.

If a conditioned reflex is not occasionally reinforced by the unconditioned (natural) stimulus it dies out—fails to produce saliva (p. 8). This phenomenon of experimental extinction has been demonstrated in children and adults as well as in animals. Razran conditioned the salivary flow of children and adults to a metronome and various visual stimuli, and later extinguished the CR by presenting the stimuli without reinforcement (mints, pretzels, and the like). The salivary flow was measured by placing a roll of cotton in the subject's mouth and, after a minute, comparing the weight of the damp cotton with the original weight of the dry cotton. Formation and extinction of conditioned reflexes was most readily accomplished when the subjects were not fully informed or were ignorant of the purpose of the experiment. Hilgard and Campbell were able to condition and then extinguish the wink reflex (lid closure) to a light stimulus. The unconditioned stimulus to lid closure was a puff of air directed into the eye.

Extinguished responses are temporarily suppressed or held in abeyance—they are not completely lost or forgotten. When a lapse of time follows extinction there will usually be some "spontaneous recovery" of the response to the conditioned stimulus. This was ascribed by Pavlov to a loss of inhibition during the interval between trials. Recovery becomes less and less, however, with successive unreinforced trials until the response to the conditioned stimulus no longer takes place. Reinforcement is then necessary to reinstate the reflex.

An interesting parallel can be drawn between extinction and forgetting. Both represent a loss of response or a decrement in available responses. But the difference is more marked than the similarity. There is good evidence that, when the retention of a lesson is measured after a time-interval, the memory loss which appears is occasioned primarily by (a) interference and confusion with other learning, and by (b) a weakening of association owing to a lack of exercise. Review and recitation are necessary in order to revive forgotten material. Extinction is not forgetting in this sense. The conditioned reflex does not die out with lack of exercise nor is it weakened with the passing of time. While there may be some loss due to interference, in the main the conditioned reflex drops out because it is extinguished—exercised without reinforcement. It can be reinstated when bolstered up by the primary (unconditioned) stimulus and sometimes it reappears spontaneously.

7. THE CONDITIONED REFLEX AND EXPERIMENTAL NEUROSIS

One of the outstanding findings in Pavlov's experimental work on differentiation is the importance of inhibition in the life of the dog. Without his "braking power" the dog would be unable to choose or select. He

would respond hit or miss to every stimulation, unimportant as well as important; in short, his existence would be a riotous confusion of events.

What happens when the dog loses its ability to discriminate is shown most strikingly in cases of "experimental neurosis" produced by Pavlov and his associates. Krestovnikova, one of Pavlov's students, conditioned a dog so that saliva was secreted whenever a circle of light was thrown upon a dark screen. As in previous experiments, the CR was built up by giving the dog food whenever the circle appeared and continuing this combination, food plus circle, until at last saliva flowed to the circle alone. Along with the circle and randomly interspersed with it, the dog was shown a small ellipse again and again, no food ever being given along with it. Eventually, the point was reached where saliva flowed at the sight of the circle, but never at the sight of the ellipse. To test the dog's discriminative ability, he was now shown larger and more circular ellipses, always without food, until finally he could distinguish—as shown by the appearance of saliva—between the circle and an ellipse whose axes were in the relation 7:8. This would seem to be a sufficient feat in itself, but the experimenter was not satisfied, and attempted to have her dog distinguish between the circle and an ellipse whose axes were as 8:9—a figure scarcely different from a perfect circle. This task proved to be too much for the dog's discriminatory ability. Saliva flowed first at sight of the ellipse, then at the circle, then at the sight of either or both without distinction. The dog began to whine, barked fiercely at the screen, tore at his restraining apparatus with his teeth, and attempted to jump down from the table. Thereafter this dog was useless as an experimental animal. Saliva would flow at sight of the experimenter, or at sight of the experimental room, or at almost any stimulus. Apparently what had happened was an almost complete collapse of the dog's discriminative ability, brought about, according to Pavlov, by too great strain being placed upon the "inhibitory brake." When finally the brake gave way, response became general and riotous.

Experimental neurosis has been produced in other animals than the dog. For example, nervous and disorganized behavior appeared in a sheep when (a) the difference between the positive and negative stimuli was *decreased* (as in Pavlov's experiment), and (b) when the *number* of presentations was increased (with shifts back and forth from the positive to the negative stimulus) without decreasing the difference between the stimuli. Electric shock was the unconditioned stimulus and a "defense reaction"—withdrawal by flexion of the foreleg—the reflexive response. Neurotic behavior has also been produced experimentally by the conditioned reflex technique in pigs and in rats. Apparently, the essential element in such behavior when produced by the conditioned reflex method is the clash between mutually incompatible responses. Faced by a dilemma,

the animal either reacts violently and without discrimination or refuses to act at all.

A loss of discriminative ability in human beings, as in animals, is often associated with nervous diseases or neuroses. The neurotic individual cannot discriminate between really dangerous and really harmless objects. Hence he is afraid of cats or dark rooms or crossing bridges, or of a thousand and one other things intrinsically harmless. Nor can he choose between really important and really unimportant tasks, and so is impelled to perform useless acts such as washing his hands ten times a day, touching every other lamp-post, going up the steps two at a time, and the like. Possibly a considerable share of the fatigue characteristic of neurasthenia comes from the large amount of wasted motion indulged in.

To be sure, even in normals, conflict and confusion (neurosis) may arise if undue strain is placed upon the discriminatory ability. A little child is allowed to shout and play roughly at one time (is even encouraged and laughed at) and is punished for doing the same thing on another occasion because the mother is tired or "company" is present. Again a child is punished for "lying" but hears his mother tell "social lies" over the 'phone—say that she's ill or is not in and hence is unable to see an unwanted caller. Many children cannot see the difference between behavior (ostensibly the same) which is sometimes right and sometimes wrong. Unable to fathom such inconsistencies, they are faced by a dilemma as real as that which confronted Pavlov's dog when it was unable to distinguish the ellipse from the circle.

Conflicts due to the interference of competing stimuli often arise in the classroom. Teach a child French and Spanish at the same time, or alternative methods of subtracting or dividing, and, for a time at least, efficiency in both tasks will be impaired. Although relatively mild, such conflicts are emotionally disturbing and are destructive of morale. Intelligent behavior in the child depends upon a judicious choice and selection from among the many competing stimuli which assail his sense organs. It is the task of the teacher to aid discrimination by emphasizing those stimuli to which responses should be made, thus preventing needless and wasteful conflicts.

8. THE CONDITIONED REFLEX AND SLEEP

Another interesting finding was Pavlov's discovery that sleep is closely linked up with the phenomena of inhibition which we encountered in the experiments on discrimination and the trace reflex. It will be remembered that in the trace reflex the conditioned salivary flow did not appear until some time after the substitute stimulus had been given. This period

depended upon the original time-interval between the substitute stimulus and the natural stimulus. Whenever the delay was one-half minute or more, a peculiar thing happened: the dog became drowsy and often went sound asleep. This even happened in delayed reflexes when the conditioned stimulus was continued through the period intervening between the substitute stimulus and natural stimulus; for example, the dog would fall asleep with the bell or buzzer going loud enough, apparently, to keep a dozen dogs awake. Annoyed by this unforeseen happening, Pavlov and his workers tried at first to get exceptionally lively and active dogs, thinking that the sleep which came so readily might be an individual peculiarity of certain dogs. Even the most active dogs, however, regularly went to sleep, indicating that some general rather than individual condition was at work. What is the nature of the sleep which occurs in these experiments? It must, said Pavlov, be a matter of general cortical inhibition analogous to, but greater than, the local inhibitory or brake effect observed in other experiments. During the waiting period between substitute stimulus and conditioned reflex, the brain and nervous system are actively inhibiting or holding in check the salivary flow. That this must be true is shown by the sudden appearance of the reflex when any strong stimulus comes in to lift the brake. When the dog's salivary flow has been positively conditioned to a note of 800 vibrations and negatively conditioned to a note of 812 vibrations, clearly there must be a delicate balance between the excitation and inhibition of the glandular (motor) response.⁵

Pavlov's theory of sleep ran as follows. Inhibition is at first evidently a matter of the local negative stimulation of a small sector in the dog's brain. But if this focus of negative stimulation is continued for a fairly long period of time (which happens in the trace reflex) the inhibition gradually spreads over the entire cortex, the dog becomes more and more drowsy, and finally goes to sleep. Normal sleep in man is readily explained after the same fashion as this experimentally induced sleep in the dog. The natural stimulus for sleep is fatigue, which induces first drowsiness (local inhibition) and then sleep (general inhibition). When the brake is lifted—the inhibition removed from the cortex—the sleeper awakens. It is interesting to note how readily sleep as a response is itself conditioned. Ordinarily sleep follows the natural stimulus, fatigue, with which is associated (usually) a comfortable bed, darkness, a certain hour, and quiet. These latter elements in the total situation often become such effective substitute stimuli for sleep that many individuals will fall asleep when in their customary situation, even though not especially fatigued. Again, for many people these associated stimuli become so important that they can-

⁵ We have already described in Section 7 what happens when too much strain is placed on the "brake."

not sleep without them, no matter how tired they may be. Insomnia, according to this view, might be thought of as occasioned by stimuli to wakefulness which repress or inhibit the stimuli to sleep; or by the absence of customary stimuli to sleep. Worry, excitement, fear, and other emotional conditions may readily act as negative stimuli to sleep.

9. THE CONDITIONED REFLEX IN ANIMALS AND IN CHILDREN

Pavlov's success in conditioning the salivary reflex of the dog stimulated many investigators, particularly in America, to undertake experiments with the conditioned reflex. Among reflexes which have been studied by Pavlov's methods are the wink, the pupillary, the flexion (hand or foot withdrawal), and the galvanic skin response. Not all of the work upon conditioning has followed Pavlov's methods. A distinction, in fact, should be made between the classical or Pavlovian experiments and those which involve what has been called instrumental conditioning. The two are fundamentally the same in that the goal in both is the formation of new associations between stimuli and responses, but they differ markedly in procedure. Typically, in a Pavlovian experiment (p. 4), two stimuli are presented repeatedly until one of them (a bell, say) is able to elicit a response (saliva flow) originally called out *only* by the other (food). In instrumental conditioning the setup is more flexible. The reinforcement (reward, escape) comes *after* the response has been made and is contingent upon it. For example, let us suppose that when a rat accidentally steps upon a disk in the floor of its cage a door is opened and food is available. Thereafter, the disk is a visual stimulus which provokes stepping upon it—a response now strongly reinforced by food. The procedure here is instrumental in the sense that the CR itself brings the reinforcement.

Conditioning experiments have been attempted with organisms all the way from protozoa to man. Many of these studies are examples of instrumental rather than of Pavlovian conditioning. All are instructive as showing the possibilities of measuring and comparing the learning processes in the lower animals. Although many of these organisms possess few activities which can be described as mental, we find, nevertheless, that they can form associations. To illustrate, in studying the behavior of the snail, Thompson found that the pressure of a piece of lettuce on the mouth of the snail provoked a chewing movement which did not appear when the pressure was applied to the foot. Here, he thought, is a chance to discover whether the snail is capable of making connections between pressure on the foot and chewing. To test this possibility the snail's mouth was repeatedly stimulated with lettuce, causing a chewing movement, and simultaneously with this pressure was applied to the foot. Continued repetition

of these paired stimuli led to the formation of a CR: whenever the foot was pressed the snail began to chew. This learned response was actually retained for four days—not a bad memory feat for a snail.

In another series of experiments, a rat obtained a pellet of food by depressing a bar which released food into a cup. Here the CR itself produced the unconditioned stimulus (food) so that the procedure is instrumental rather than Pavlovian. Extinction of this CR in rats and its reappearance through reinforcement have been carefully investigated, and curves describing these phenomena plotted. If a dog's foot is pricked with a pin or given an electric shock, it will be pulled away quickly and involuntarily. By using the electric shock as an unconditioned stimulus, this "withdrawal reflex" of the dog may be so conditioned that sounding a note will cause the foot to pull back.

The CR technique has proved useful in the study of sensory discrimination in various animals. A carefully devised series of experiments on the formation of CR's in fish has been reported from England, with food as the natural or unconditioned stimulus. Fish were conditioned to a number of substitute stimuli such as a slight rise in temperature of the water (as small as $.4^{\circ}\text{C}$); changes in the salinity of the surrounding water; and to auditory (probably vibratory) and visual stimuli. By positively conditioning the saliva flow of a dog to a given tone and negatively conditioning it to other tones, auditory acuity may be accurately determined (p. 9). Auditory acuity in monkeys has been shown by the conditioned response technique to be fairly comparable to that of man.

Many cases of animal learning are readily explainable in terms of instrumental conditioning. Performing animals, dogs, horses, elephants, and the like, are usually responding to substitute stimuli such as the trainer's voice, a movement of his head, or a movement of his hands, all of which have frequently been reinforced in the past. Illustrations showing how the CR is employed in the training of animals will be found in Chapter 3, pp. 54–55.

As little children are able to introspect little if any better than animals, the CR method is especially well adapted to the study of their responses. Krasnogorski, one of Pavlov's students, was the first investigator to study learning in young children by the CR method. He conditioned the salivary reflex in several young children to a variety of stimuli, with a fair degree of success—the sight of food, a bell, sound of a reed pipe, a slight scratch on the skin. As an indication of the amount of saliva secreted, Krasnogorski noted the number of mouth-openings and swallowing movements made by the child. He states that conditioned reflexes can be built up as early as the first year of life; and that almost any stimulus can serve as a substitute stimulus for a reflex motor or secretory activity. An important discovery was that conditioned reflexes break down more readily in normal than in

abnormal and feeble-minded children. Apparently the feeble-minded child is less flexible and more mechanical in his associations than is the normal.

In America, Mateer, Marquis, and Razran have extended and verified much of Krasnogorski's work. Mateer dealt with more than fifty normal children from about twelve to eighty-nine months of age, and with a half dozen or so feeble-minded children. She studied the formation of the conditioned reflex, its retention over a period, its dying-out or extinction (unlearning), and the ease with which conditioned responses could be re-established. Mateer worked with the salivary reflex, noting the number of swallowings, as did Krasnogorski, but her experiments were much better controlled than his. Conditioned reflexes in normal children were learned and extinguished in about one half the number of trials necessary for the feeble-minded. Mateer also found a substantial relationship between the ease of conditioned reflex formation and mental ability as measured by tests. This author regards the conditioned reflex method as especially valuable in work with young children. It provides a direct measure of learning ability and is a useful supplement to other clinical methods.

10. THE CONDITIONED REFLEX IN ADULTS

The extent to which the reflexes of the human adult may be conditioned has been attacked experimentally by many psychologists. Knowledge of what is being attempted, an indifferent and often an unfavorable attitude on the part of the subject, and the exercise of more voluntary control over his responses, renders the conditioning of reflexes in adults less dependable and more difficult than is the case with animals and children. Cason conditioned the pupillary reflex (change in size of the pupillary aperture) by ringing a bell each time a beam of light was thrown in the subject's eye. After 400-odd joint stimulations of bell plus light, the pupil dilated at the sound of the bell alone. In one subject, the pupil was even conditioned to contract to the bell—a result which shows the tremendous possibilities for connection forming in the human nervous system. The wink reflex (lid closure) is readily conditioned as is also the galvanic skin reflex. Experiments with the knee jerk have given contradictory results. This reflex (a sudden jerking forward of the leg when the patella tendon is struck a sharp blow) can probably be conditioned to sound, though this is accomplished with difficulty and with much irregularity of results. Several experimenters have succeeded in conditioning hand withdrawal to sound (by using shock as the unconditioned stimulus), but it seems probable in many cases that the response is at least partly voluntary. Watson was one of the first to condition withdrawal of the foot as well as withdrawal of the

finger. An electric shock was the unconditioned stimulus, while a bell, a buzzer, and a bright light were among the substitute stimuli.

11. THE CONDITIONED REFLEX AS A PRINCIPLE OF LEARNING

There is perhaps no subject in experimental psychology upon which more time and effort have been expended than upon the conditioned reflex. The acquisition of conditioned reflexes by animals, children, and adults; the relative ease of conditioning various reflexes; the stability of conditioned reflexes, their extinction and reappearance; the relation of school learning to ease of formation of conditioned reflexes; the qualitative and quantitative distinctions between natural and conditioned reflexes—these and other aspects of the phenomena of conditioning have been subjected to experimental attack. Pavlov's method appealed to the early experimentalists because it offered the possibility that complex activities could be thought of as composed of simple links forged into a more intricate pattern. Furthermore, the conditioned reflex seemed to substitute an exact and measurable relationship for explanations in terms of trial-and-error and "satisfyingness." Many psychologists hoped—and the strict objectivists believed—that the conditioned reflex would prove to be the unit or element out of which all habits are built.

Faith in the Pavlovian conditioned reflex as the fundamental unit of learning is now largely a matter of historical interest. Learning is not, as Pavlov thought, simply a "long chain of conditioned reflexes," but is a coöperative and organizing enterprise in which the whole organism takes part. But objection to the classical conditioned reflex as the unit of learning does not imply that the concept of conditioning is not useful, in fact well-nigh indispensable, to learning theory. Psychologists have leaned heavily upon instrumental conditioning as an explanatory principle in learning and perception. The CR has been used to explain the origin of gestures, language, and various group phenomena, as well as personal and emotional adjustments. Methods of reconditioning children and adults in order to remove bad habits, personality defects, and various maladjustments are based essentially upon the concepts of reinforcement and extinction.

In summary, it can be said that Pavlov's contribution to experimental psychology lies in his having provided a precise technique for studying how new responses are acquired, how responses are lost, and how revived. No learning theory can fail to take account of Pavlov's work: in fact, none can exist without it.

BIBLIOGRAPHY

1. Allport, F. H. *Social psychology*. Boston: Houghton Mifflin Company, 1924.
2. Anderson, O. D., & Liddell, H. S. Observations on experimental neurosis in sheep. *Arch. Neurol. Psychiat.*, 1935, 34, 330-354.
3. Anrep, G. V. Pitch discrimination in the dog. *J. Physiol.*, 1920, 53, 367-385.
4. Bull, H. O. Studies on conditioned responses in fishes. Part I. *J. Marine Biol. Ass. United Kingdom*, 1928, 15, 485-533.
5. Cason, H. The conditioned pupillary reaction. *J. exp. Psychol.*, 1922, 5, 108-146.
6. Dashiell, J. F. *Fundamentals of general psychology*. Boston: Houghton Mifflin Company, 1949.
7. Dodge, R. Problems of human variability. *Science*, 1924, 59, 263-270.
8. Frolov, J. P. *Pavlov and his school: the theory of conditioned reflexes*. New York: Oxford University Press, 1937.
9. Guthrie, E. R. *The psychology of learning*. New York: Harper & Brothers, 1935.
10. Hilgard, E. R. The relationship between the conditioned response and conventional learning experiments. *Psychol. Bull.*, 1937, 34, 61-102.
11. Hilgard, E. R., & Campbell, A. A. The course of acquisition and retention of conditioned eyelid responses in man. *J. exp. Psychol.*, 1936, 19, 227-247.
12. Hilgard, E. R., & Marquis, D. G. *Conditioning and learning*. New York: Appleton-Century-Crofts, Inc., 1940.
13. Hilgard, E. R., & Marquis, D. G. Acquisition, extinction, and retention of conditioned lid responses to light in dogs. *J. comp. Psychol.*, 1935, 19, 29-58.
14. Hollingworth, H. L. *Abnormal psychology*. New York: The Ronald Press Company, 1930.
15. Humphrey, G. Extinction and negative adaptation. *Psychol. Rev.*, 1930, 37, 361-363.
16. James, W. *Principles of psychology*. New York: Henry Holt and Company, 1890. Two volumes.
17. Lashley, K. S. Nervous mechanisms in learning. Chap. 10 in *Handbook of general experimental psychology*. Worcester, Massachusetts: Clark University Press, 1934.
18. Maier, N. R. F. *Studies of abnormal behavior in the rat*. New York: Harper & Brothers, 1939.
19. Marquis, D. P. Can conditioned responses be established in the newborn infant? *J. genet. Psychol.*, 1931, 39, 479-492.
20. Pavlov, I. P. *Conditioned reflexes*. London: Oxford University Press, 1927.
21. Pavlov, I. P. *Lectures on conditioned reflexes*. New York: International Publishers, 1928.
22. Razran, G. H. S. Conditioned responses in children. *Arch. Psychol.*, New York, 1933, 23, No. 148.
23. Razran, G. H. S. Conditioned responses: an experimental study and a theoretical analysis. *Arch. Psychol.*, New York, 1935, 28, No. 191.
24. Razran, G. H. S. Conditioned responses; a classified bibliography. *Psychol. Bull.*, 1937, 34, 191-256.

25. Shaffer, L. F. The psychology of adjustment. Boston: Houghton Mifflin Company, 1936.
26. Skinner, B. F. On the rate of extinction of a conditioned reflex. *J. gen. Psychol.*, 1933, 8, 114-129.
27. Thompson, E. L. An analysis of the learning process in the snail, *Physa gyrina* Say. *Behav. Monog.*, 1916, 3, No. 14.
28. Thorndike, E. L. The fundamentals of learning. New York: Teachers College, Columbia University, 1932.
29. Watson, J. B. The place of the conditioned reflex in psychology. *Psychol. Bull.*, 1916, 23, 89-116.
30. Wendt, G. R. Auditory acuity of monkeys. *Comp. Psychol. Monog.*, 1934, 10, No. 3.
31. Wolffe, H. M. Conditioning as a function of the interval between the conditioned and the original stimulus. *J. gen. Psychol.*, 1932, 7, 80-103.
32. Woodworth, R. S., & Marquis, D. G., *Psychology*, 5th ed., New York: Henry Holt and Company, 1947.
33. Woodworth, R. S. *Experimental psychology*. New York: Henry Holt and Company, 1938.
34. Yerkes, R. M., & Morgulis, S. The method of Pavlov in animal psychology. *Psychol. Bull.*, 1909, 6, 257-273.

2

FRANZ'S AND LASHLEY'S STUDIES OF THE RÔLE OF THE BRAIN IN LEARNING

I. THE BRAIN IN LEARNED ACTIVITIES

THE PIONEER work of Shepherd I. Franz and Karl S. Lashley is an excellent example of the physiological approach to problems of learning and habit-formation. Both of these scientists were trained in physiology as well as in psychology. Franz received the doctor's degree in psychology from Columbia University in 1899, and in 1902 began a series of studies on the rôle of the brain in learning, using cats and monkeys as experimental subjects. For many years he was psychologist at the Government Hospital for the Insane in Washington, where he did striking work in the reëducation of individuals who had sustained brain injury. Lashley attained the Ph.D. degree at the Johns Hopkins University in 1914, and was for a time a student of Franz. He is the most distinguished American worker in the field of physiological psychology.

The task which Franz and Lashley set for themselves was to determine *what* regions of the brain and *how much* of the brain functions in specific learned acts. Their method was to compare the learning ability of animals before and after destruction of certain brain areas by operation; and also to compare the learning ability of "operated" animals with that of normal animals. A group of rats, for example, is trained to thread a maze, the number of trials and the time being recorded; parts of the brain tissue are then removed and after recovery from the operation the animals are re-tested to see to what extent the loss of brain structure will affect their behavior. Or again, animals whose brains have been partially destroyed are trained to escape from a problem box, and their records are compared with those of normal animals.

Operations involving different brain areas have been performed upon rats, cats, and monkeys, and the effect of brain loss upon the acquisition of both sensory and motor habits of different degrees of complexity has been investigated. Experiments like these have much significance for human learning. The human brain cannot be experimented with directly and it

1697

is only in the event of disease or wounds that brain injury can be correlated with loss of mental function (pp. 37-38). In pathological conditions of the brain, however, complications arising from bodily disease or injury must be seriously considered. Such disturbances may be so drastic that a definite conclusion as to the specific effects of the brain injury itself is often hard to reach. This situation does not arise in operative work with animals; parts of the cerebral cortex in a rat, say, can be experimentally destroyed and the resulting effects upon learning behavior directly noted. Moreover, experimentation with animals permits us to observe the rôle of the brain in behavior less complex and more easily controlled than in man.

Experimental findings from the higher animals, e.g., the apes, can be carried over almost unaltered to human learning, and even in the lower species problems in the physiology of learning are fundamentally the same as in man. In a careful review of the literature on cerebral function, Lashley concludes that while there is undoubtedly greater specialization as we go up the evolutionary scale, the relation of brain areas to learning does not differ markedly in man from that in other mammals.

2. EARLY STUDIES OF BRAIN FUNCTION AND LEARNING

Before considering the experimental work on brain function in animals, let us get as clear a notion as possible of the topography of the brain itself. The brain of the rat, cat, or monkey is much simpler—but is basically similar—to that of man. Hence we can locate the positions of the various structures in the animal brain by reference to Figure 5, which gives a side view of the left hemisphere of the human brain. The frontal "lobes" or areas lie just behind the forehead, while behind the frontals and separated from them by a great fissure are the parietal lobes. The parietals lie under the "roof" of the head. The occipital lobes, concerned primarily with seeing, lie at the back of the head; the temporal lobes, concerned in part with hearing, are just above the temples. The adjoining areas of the parietal and temporal lobes are the great sensory synthesizing and associative areas of the brain, while the frontal lobes constitute the motor organizing and motor discharging centers. The frontal lobes in man are also believed to be concerned with the management and restraint of the more primitive impulses.

In his first studies Franz investigated the rôle of the frontal lobes in the retention of associations formed in learning. His experiments were carried out with cats and monkeys. The cats were first trained to escape from boxes quite similar to those used by Thorndike in his studies of animal learning (see p. 43). These boxes were about twelve inches high, fifteen inches wide, and twenty inches long. The bottom, back, and sides were solid, the

6.01.2009
13815



front and top covered with three-quarter inch slats. The door was six inches wide by eight inches high, and was hinged at the bottom so that it would fall forward when the latch or other mechanism holding it was released. Three boxes were used. Two could be opened by pushing or knocking against a cord attached in different ways to the door; the third by pressing a button up or down. This third box is shown in Figure 6.

The stimulus to activity was food—meat, fish, or milk—placed just outside of the box. Given a training period of about a week, a normal and hungry cat can learn to escape from boxes such as these in from two to six seconds. After the cat had learned the habit to this high point of efficiency,

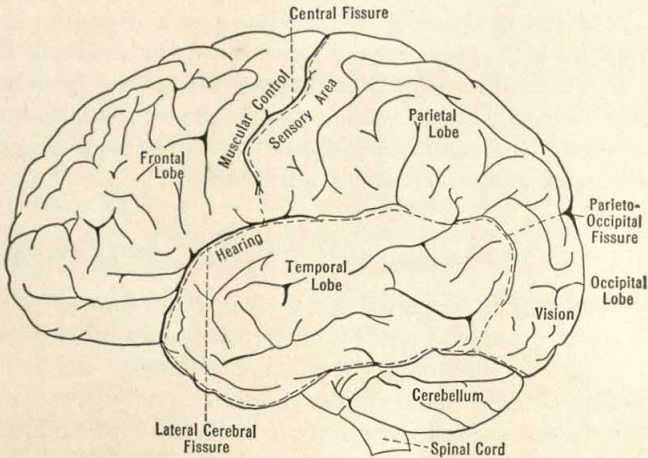


FIG. 5

THE LEFT ASPECT OF THE CEREBRAL CORTEX IN MAN (SCHEMATIC)

Franz operated under anesthesia, cutting away both frontal lobes from the rest of the brain structure without, however, removing the severed parts from the skull. After the animal had recovered from the anesthetic and from the shock following the operation—which usually required two to three days—it was retested in the box from which it had previously learned to escape. In all three cats used in this experiment the habit of escape, though repeatedly tested for, failed to reappear. But the cats still made impulsive efforts to escape; also they retained old and much-used habits, such as purring when petted, rubbing against the experimenter's legs, and responding to being called. When they were held before a cage of mice, normal emotional responses were exhibited, such as increased heartbeat, following the mice rapidly with the eyes, and attempts to jump. Only the *recently acquired habit* of opening the box seemed to be lost. Franz was careful not to use the results from any animal in which postoperative in-



KARL S. LASHLEY

1890-



fection had set in, and in every case the cat was later killed and its brain examined to make sure that the frontals had actually been separated from the rest of the brain.

The loss of learned responses following extirpation of the frontal lobes holds for monkeys as well as for cats. In his experiments with monkeys, Franz used small animals of the short-tail (rhesus) and the ring-tail variety. Two tasks were arranged, the one requiring the opening of a food box, and the other the obtaining of food by completing a so-called hurdle task. The problem set by the food box was to open the door of a small food com-

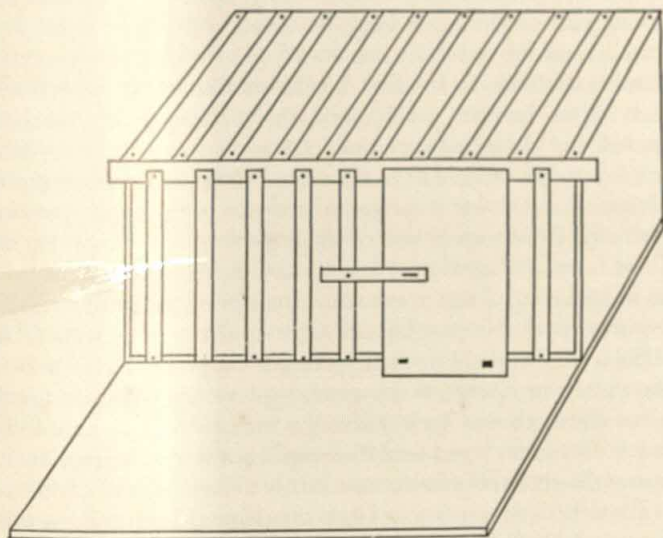


FIG. 6

PUZZLE BOX USED BY FRANZ IN LEARNING EXPERIMENTS WITH CATS

partment attached to the cage in which the animal lived. The door of the food box was held by a button much like that in the door of the box in Figure 6. In the hurdle problem the animal had to perform several acts in order, going from one bar to another, crawling through a hole, and so forth, before reaching the food. All of the animals were hungry when placed in the test situations so that active and vigorous behavior was readily obtained.

Six monkeys were used in this experiment. As soon as they had thoroughly learned the habits, parts of their frontal lobes were removed, with the same precautions employed as in the cat experiments. When the monkeys were retested after recovery from the operation, all were found to have lost the trick of opening the food box or of going through the more

complex activities of the hurdle. But as in the case of the cats, habits of long standing, such as eating out of the experimenter's hand and jumping upon his shoulder, remained. The monkeys were described as normally active and apparently normal emotionally, screaming and chattering as before the operation.

The conclusion reached in these experiments, namely, that newly formed associations and habits are primarily dependent upon activity of the frontal lobes, encountered several objections. In the first place, it was argued that the shock of the operation (surgical shock) might alone be sufficient to break up the new associations; and secondly, that mere excision of brain tissue in *any* part of the brain (not the frontal lobes particularly) might explain the loss of the habit. To answer these criticisms, Franz performed two control operations. In the first, a cat was trained as before to escape from a box. After the trick was learned, the cat was anesthetized and the skull opened, but no incision of brain tissue was made. The wound was simply closed and bandaged as in the other operations. Twenty-four hours after the operation, this cat was retested and was found to have retained the habit perfectly. Other repetitions of this experiment indicate that surgical shock alone is not sufficient to obliterate new associations.

In the second control test, a cat which had been trained to escape from two boxes was operated upon, but this time a portion of the *parietal lobe* instead of the frontal was cut away from the underlying parts of brain tissue. Three days after the operation this animal got out of the box in a little over seven seconds, which was about twice the time required before the operation. Shortly thereafter it reduced its escape time to the former level, showing, apparently, that the associations involved in the habits had not been actually disturbed. Seven weeks later, *one* frontal lobe only in this same cat was excised. Four days after this operation, though extremely slow, the animal was able to do the trick. In another cat which had learned perfectly how to get out of the box, *both* parietal lobes were destroyed. Three days after the operation, this animal performed the trick as well as before. These control experiments indicate clearly (1) that only when *both frontals* are injured are the recently acquired associations lost; and (2) that the loss of the habit cannot be explained as due either to surgical shock resulting from the operation or to mere cutting of the brain tissue.

3. THE VICARIOUS RÔLE OF BRAIN STRUCTURES: RESTITUTION OF FUNCTION

In all of Franz's experiments described so far, habits and associations lost as the result of frontal injury were of quite recent origin; old and well-established habits persisted. To see whether this would hold true for

highly trained escape habits, in a second series of experiments Franz had four monkeys practise over a much longer period of time until the escape habit was firmly established—almost second nature, so to speak. The frontal lobes of these highly trained animals were then destroyed, and after recovery it was found that the escape habit functioned almost, if not quite, as perfectly as before the operation. In another series of experiments with several cats, the frontals were severed from the rest of the brain *before* the animals had learned to escape from any of the boxes. When these animals were trained later, it was found that they could learn the habits as well as normal cats.

These experiments are interesting and suggestive—they indicate that in an emergency one part of the brain can take over the function of another part vicariously, or by substitution. It seems evident from Franz's results that in cats and monkeys recently formed complex motor habits are normally carried by the frontal lobes since the habits are lost when the frontals are removed. But if the habits are well ingrained, i.e., well mechanized, they persist even when parts of the frontals are lost, being carried on apparently through the agency of other parts of the brain. That these other centers *may* provide a basis for the habit in the first place is indicated by those experiments in which cats whose frontal lobes had been excised *before* training still learned the habit. These findings are opposed to the old idea of specific "learning centers" in the brain which control definite activities. Such areas, if they actually exist, must be considerably more diffuse, and more loosely organized, than was originally believed.

Several cases of restitution of function taken from Franz's work with human beings are instructive in this connection. These cases are concerned with the reëducation of patients whose brains had been injured by disease or wounds. An aphasic, fifty-seven years old, had recently suffered a paralytic stroke.¹ His right side was paralyzed, and his language was jumbled and often unintelligible. Many irrelevant words were employed and objects were often named incorrectly. The attempt was made to retrain this patient by having him relearn (1) the names of ten familiar colors; (2) a short stanza of poetry; and (3) the Lord's Prayer. Beginning with an accuracy score of 44 per cent right in naming colors, after three months, in which the trials varied in number from day to day and were sometimes omitted, this man increased his accuracy score to 96 per cent right. Comparable results were obtained in relearning numbers. An interesting finding here was that the number 5 (probably because of its common use) was

¹ *Aphasia* is loss of ability to understand spoken or written language (sensory type); or loss of ability to say the correct word to express one's thoughts (motor type). Aphasia is, in general, the result of injury to the associative areas of the cortex, most often, probably, in the frontal and temporal lobes.

most often used spontaneously. In trying to learn poetry—the first stanza of *The Village Blacksmith* was the selection used—the patient read aloud the first one or two lines five times and then tried to recite. Repetition was continued until reproduction was perfect, whereupon the next few lines were attempted, and so on. After five days of work the patient was able to repeat six lines of the selection sufficiently well to indicate that he was gradually learning it. The Lord's Prayer, formerly well known to this patient, was correctly given after twenty-five readings spaced over three days. A post-mortem examination of this man's brain (he died of pneumonia shortly after the experiments described) showed considerable destruction in the lower part of the cortex, i.e., in the temporal lobes. Franz points out that his slow but steady reëducation suggests the possibility of the reëstablishment of old brain connections or the opening-up of new ones. This seems reasonable in the light of the experimental work on animals which has been described.

A second case illustrates the powerfully stimulating effect of positive suggestion. In this instance the patient was an ex-soldier, a young man who had been struck by a high-explosive shell. The result of his injuries was paralysis of face, arm, and leg on one side. This condition had persisted for nine months. This man entered the examination room hobbling on a cane. At once he was ordered sharply by the examiner to put down his cane and take a seat in a chair some twelve feet away. Although he insisted at first that he couldn't walk without a cane, upon being told that the examiner was quite sure that he could, he laid the cane aside and awkwardly, but unassisted, walked to the chair. Here the authority and the prestige of the examiner proved to be a sufficiently emphatic stimulus to reinstate a partly lost function. Either the old nerve connections conducted under "pressure" or other and newer connections substituted for them.

The effect of a strong incentive or an emotional stimulus in making a patient "forget" his disabilities is well illustrated in the case of another partially paralyzed man who had walked with a cane for nineteen years. During a baseball game this individual hit the ball, and in great excitement, without pausing to get his cane, ran quickly to first base, beating out the throw. It is significant that in spite of this achievement he then demanded his cane, saying that he couldn't walk without it. In many cases of this sort it seems likely that old pathways in the brain and nervous system are still partially intact, or that others are able to substitute for them. The patient's attitude is often the controlling factor in stirring these connections into life. If the suggestions are strong enough, or if the patient's interests can be sufficiently aroused, the strengthened motivation may lead to recovery of the lost skills. In some instances, strong emotional incentives—anger or fear, say—may be necessary in order to provoke into activity long dormant

functions. Apparently reëducation depends more upon the emotional impetus and the will-to-do than upon the degree of physical handicap to be overcome.

4. THE BRAIN IN BRIGHTNESS DISCRIMINATION

Lashley's work, like that of Franz, was directed toward discovering the effect of brain destruction upon specific learned acts. In one series of experiments a group of white rats was trained in a very simple sensory habit of brightness discrimination. The problem box used (Yerkes Discrimination Box, shown in Figure 7) contains two alleyways, either of which can

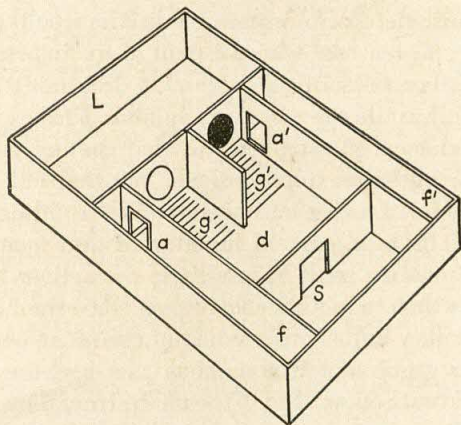


FIG. 7

BOX USED FOR TRAINING IN BRIGHTNESS DISCRIMINATION

The animal is started at S and passes to d, where a choice is offered of an illuminated and a dark alley. These alleys lead across electric grills (g, g') through trapdoors of light celluloid (a, a') to the food compartments (f, f'). In training, the door on the dark side is locked and the grill charged. The illumination is shifted irregularly from side to side by a movable screen in the light-box (L). When in use the entire box is covered except for an observation hood above d. (From Crafts, Schneirla, Robinson, and Gilbert, *Recent Experiments in Psychology*, 2nd ed., after Lashley, 1929. Copyright, 1950. McGraw-Hill Book Company, Inc.)

be illuminated by a transparent screen placed at the far end of the box. When the rat takes the lighted side, it is admitted to the food compartment through a small door; when it takes the dark side it receives no food, but instead gets a shock through an electric grill placed on the floor of the alley. The light can be shifted from the left to the right alley irregularly so that the animals cannot form a place habit. About 100 trials are required for a rat to fix the association between the illuminated alley and food. Training was continued at the rate of ten trials per day until thirty consecutive

errorless trials were obtained. After one group of animals had learned the habit perfectly, varying amounts of the occipital lobes of both cerebral hemispheres were destroyed by operation. After recovery from the operation the rats were tested for retention. *All* of these animals were found to have lost the brightness discrimination habit. But with practice they were able to reacquire it, the number of trials needed being roughly proportional to the *extent*, but independent of the *location*, of occipital lobe destruction. The correlation between percentage of cortex destroyed and retardation in learning the brightness discrimination habit was .45.

Another group of animals untrained in brightness discrimination was first subjected to brain operations which involved different *one thirds* of the occipital cortex. These rats were then tested to see how long it would take them to acquire the discrimination habit. The results of these tests are quite convincing. Those rats whose brains were injured *before* opportunity had been given to learn the brightness discrimination acquired the habit just as quickly as did the uninjured animals. This experiment and the one just described seem to establish (1) that the occipital lobes are ordinarily active in brightness discrimination, but that (2) no part of these areas is necessary for it, as various parts—up to a critical limit of 50 per cent—can apparently take over the function. It also seems clear that (3) brightness discrimination must be related to the activity of the *whole* occipital lobe rather than to any specific region, since the habit is weakened by the extent of injury but is independent of the point where it occurs.

In still another group of rats trained in the brightness discrimination habit the parietal brain areas (Fig. 5) were destroyed by operation. Upon being retested, these animals, although considerably slower than before, in no case showed complete loss of the habit. This result suggests that the occipital areas were the regions originally concerned in learning this activity. We have already seen, however, that when parts of the occipital lobe were destroyed *before* training, the rats were still able to learn the brightness discrimination habit as well as normal animals. This led Lashley to propose the following query: *Is there any particular area of the brain which takes over the functions of the injured occipital parts?* To answer this question, he destroyed different parts of the occipital lobes in twelve rats and trained them until they had learned the brightness discrimination habit up to his standard. A second operation was then performed on the same rats, various regions of the remaining parts of the cortex being destroyed. Each operation injured about one third of the still intact cortex, the injuries sampling the remaining area in such a way as to cover (when all twelve rats are considered) the entire two hemispheres. In no animal was *all* of the brain cortex destroyed, as complete decortication has not as yet been successfully performed in the rat. These twice-operated animals

were then retested for retention of the brightness discrimination habit, with the result that none was found to have lost it.

This last experiment forces us to the important conclusion that apparently *any* part of the cortex is able to take over the brightness discrimination habit in the absence of those areas whose function it normally is. Since the entire cortex was not removed in any animal used in these tests, it is uncertain whether *some* cortical tissue is not necessary for the retention of the brightness discrimination habit. But after occipital injury no *specific* area is needed to carry on the habit vicariously. One possible explanation of this phenomenon is that subcortical centers² take over the functions of the injured cortex. This would seem probable in view of the widely held opinion that habits and skills acquired through cortical activity (i.e., by conscious effort) will—when thoroughly automatized—be short-circuited to subcortical levels. But Lashley's work offers no evidence in support of this view. In one experiment, rats were overtrained by 1,200 trials, that is, by more than 10 times the number necessary to establish the brightness discrimination habit. In these animals the brightness habit had surely become highly automatic. After cortical destruction, however, these rats did not relearn any more readily than rats which had received normal training. Lashley writes: "There is no indication that subcortical nuclei have taken over any part of the reaction, even sufficient to facilitate learning." If subcortical centers do not take over the function of the injured parts of the cortex, there must be reorganization or substitution in the remaining intact cortex.

These results which have been quoted from Lashley's work are quite securely established. More than 150 rats were used as experimental subjects and there were adequate controls, so that the results cannot be attributed to surgical shock or to actual blindness in the animals. Post-mortem examinations to verify the extent of the lesions were also made.

5. BRAIN FUNCTION IN MOTOR HABITS

In another series of experiments, Lashley studied the effects of brain injury upon the learning and retention of a motor habit. Animals were first trained to secure food in the problem box shown in Figure 8. In this box, called the double platform box, it was necessary for the animal to depress platforms *a* and *b* in the fixed order *a—b* in order to open the door, *d*, which led to the food compartment. These platforms were about four by two inches and were raised one inch from the floor; a slight pressure was sufficient to depress either of them. This task required more

² Subcortical centers are structures in the older and more primitive brain-stem which are overlaid by the cerebrum and cerebellum.

complex motor responses than were needed in the inclined plane box used in earlier experiments by Franz and Lashley. In the inclined plane box, the animal had to climb upon the food compartment (a small box placed within the cage) and depress one small platform in order to open the door of the food box. In the double platform box, *two* platforms instead of one had to be depressed and in the right order.

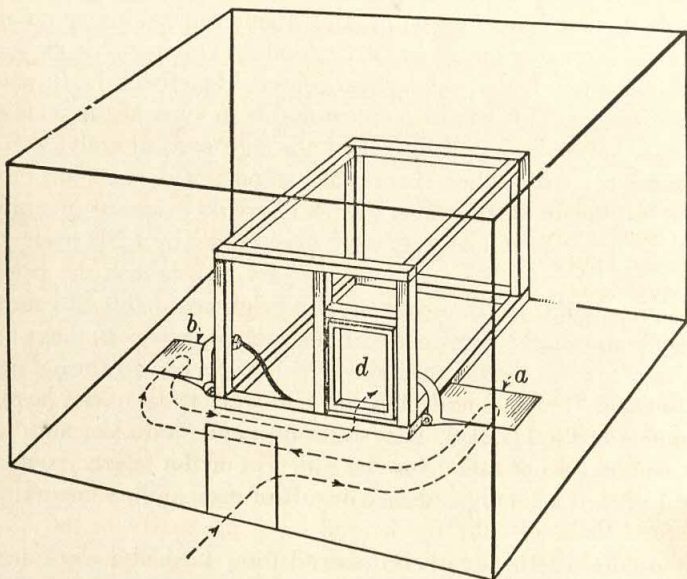


FIG. 8

DOUBLE-PLATFORM BOX USED IN EXPERIMENTS IN LEARNING WITH WHITE RATS
The door *d* is opened when the platforms *a* and *b* are successively pressed down.
(From Lashley, 1929.)

Nineteen rats were trained in the double platform box *after* about one third (from 14 to 50 per cent), of the cerebral cortex had been destroyed by operation. These operated animals learned the habit as readily as the normal (control) rats, actually requiring only about three fifths as much practice, as shown in Table I. This rather remarkable result may be partly explained by the fact that the operated rats showed somewhat less distractibility and consequently a smaller amount of random activity. Animals with brain injuries jumped over the platforms less often than normal rats, more often bumping into and eventually depressing them. The task of opening the door of the food compartment also seemed to be easier for the injured rats. Lashley found, for instance, that operated rats slightly paralyzed and with motor disturbances required about 30 per cent fewer trials than other operated rats which showed no motor defects.

In certain tasks, therefore, it appears that brain destruction actually favored quick learning, but it did so by slowing up and often drastically reducing voluntary activity.

TABLE I

MEAN NUMBER OF TRIALS AND THEIR RANGE FOR NORMAL AND OPERATED RATS IN LEARNING THE DOUBLE PLATFORM BOX

(After Lashley, 1920)

<i>Group</i>	<i>Injury</i>	<i>Mean No. Trials</i>	<i>Range</i>	<i>No. of Animals</i>	<i>Average Per Cent of Brain- Destruction</i>
1	Normal	142.6	63-204	10	0
2	One hemisphere destroyed	87.2	49-141	6	37
3	Occipital injury	63.8	45-107	4	28
4	Parietal injury	80.0	41-101	5	22
5	Frontal injury	90.0	90	2	16
6	Fronto-parietal injury	39.0	27-51	2	28
All operated rats		79.0	27-141	19	28

In subsequent experiments rats first trained in the double platform habit were subjected later on to operations which destroyed various parts of their frontal and occipital brain areas. The motor habit, though much disturbed, was not totally lost as a result of such destruction. This finding suggests that normally the frontal lobes are active in the formation and retention of motor habits. But substitution of function by other areas is immediate and often remarkably effective.

Lashley has checked and extended many of his earlier results on learning and retention in rats. Tests of the learning ability and retention of both motor and sensory habits were made with four mazes of different pattern and with the brightness discrimination box shown in Figure 7. The experimental group was made up of fifty animals, all of which were subjected to brain operations involving (in a given case) from 1.5 per cent to 81 per cent of the total surface area of the cerebrum. A few of the more significant findings in these careful and large scale experiments may be summarized briefly. In substantiation of his earlier work, Lashley found (1) that the capacity to form motor habits (for example, to learn to run mazes) is reduced by brain destruction, the reduction being independent of the place of the injury but roughly proportional to the amount of destruction. The correlations between extent of cortical destruction and retardation in learning ranged from .70 to .80. (2) The more complex the problem, the greater the effect upon learning produced by *any* given injury. (3) The habit of brightness discrimination

is not greatly affected by cerebral lesions even when a considerable amount of the cortex is destroyed. The discrimination of visual forms and patterns such as circles, squares, and the like is more affected by cortical lesions than is the brightness habit. (4) Retention of simple motor habits (learning mazes) after forty days is significantly impaired by cerebral injury. The degree of impairment depends chiefly upon the extent of the injury and the initial learning ability of the rat.

6. EQUIPOTENTIALITY AND MASS ACTION OF THE BRAIN

Perhaps the most far-reaching result of Franz's and Lashley's experimental studies of the brain is their implication that the learning and retention of sensory and motor habits cannot be explained in terms of fixed nervous pathways, definite brain structures, or specific synaptic connections. Apparently, complex learned acts draw upon more extensive brain areas and depend more upon the total organization within the cortex than was formerly thought probable (see Chapter 4). Lashley has used the term "equipotentiality" to designate the capacity of an uninjured part of the brain to take over functions lost by destruction of other areas. Equipotentiality is subject to the law of "mass action"; that is, the efficiency of performance of any complex activity is reduced in proportion to the *amount* of brain injury—the mass of cortex destroyed—but is not dependent upon the integrity of any particular region. Equipotentiality seems to hold in subcortical as well as in cortical destruction. In one experiment, Ghiselli and Brown destroyed various subcortical regions in 45 rats. These animals were inferior to normal rats in learning an eighteen-unit maze, but no particular subcortical area proved to be indispensable to maze learning.

The brain of the rat is comparatively simple and undifferentiated as compared to that of the monkey or of man. It has been argued, therefore, that Lashley's results cannot be applied to higher species. It is true that to generalize from results obtained with the rat directly to man is fraught with danger. But experimental work with the higher apes, as well as studies of human beings whose brains have been injured by war or disease, have tended to confirm Lashley's theory of equipotentiality and mass action. In a study of the frontal associative areas in the ape, Jacobsen found that the loss of previously learned responses depended upon whether the lesion involved one or both frontal lobes. Further, in a study of monkeys, Jacobsen and Taylor report that the recovery of once learned activities depends upon the amount of injury to the frontal lobes. Complete recovery followed slight lesions; very little recovery followed large lesions. Jacobsen's results indicate that while there is evidence for

equipotentiality in the ape's brain, the loss of specific memories (and the retention of others) following lesions of the frontal lobes argues for less mass action in the ape than in the rat. Pike and Chappell, in experimental studies of the effects of brain lesions in cats, report greater specialization of brain structures in cats than Lashley found in rats. These authors stress the importance of organization within the nervous system, and insist that certain structures have definite functions.

The importance of organization in restitution of function after brain injury is well illustrated in an experiment by Klüver. Monkeys were trained to discriminate between pairs of weights, sounds, visual forms and colors. First the animals learned to pull into their cage by strings *one* of a pair of boxes upon which the stimuli to be discriminated were fastened. After the monkeys had learned the trick, the occipital, frontal, and parietal areas in different animals were destroyed by operation. For several weeks following the recovery from the operation, the monkeys pulled in the boxes at random, seemingly having forgotten the discrimination habit completely. They were then carefully retrained with a pair of weights and again taught to make a comparison of the resistances offered by the two strings when first one and then the other was pulled. After this training, the animals were able immediately to make correct comparisons of the visual and other stimuli although these discriminations had apparently been completely lost. Training in the basic process of comparing enabled these animals to recover and to utilize lost or forgotten habits which depended upon it. Recovery of function after brain injury would seem, then, to depend upon successful *reorganization* of available responses by training or otherwise as much as or more than upon the *amount* of still intact cortical tissue.

Many studies of the relationship of brain injury to mental disturbances in man give evidence of equipotentiality and mass action. An interesting example is the famous "crow bar" case, reported by Harlow in 1848, in which a laborer lost a large part of the frontal areas of his brain when a crow bar was driven through his head in a blasting accident. According to accounts this man showed little intellectual impairment after recovery, but his temper became more violent and his emotional control was lessened. Evidence of this sort is not, of course, always convincing. After World War I there were many cases in which patients having lost relatively large amounts of brain tissue were reported to have recovered completely. "Complete recovery" seems often to have meant no more than walking across the room and answering a few simple questions. In animal work the same loose statements concerning "normal behavior" are often found. After a brain operation an animal's behavior may appear to be entirely normal to the casual observer, even though careful

tests show marked retardation in learning. In fact, the point should be emphasized, that *any* brain injury will affect behavior in some degree.

A striking and well-documented case showing the effect of loss of brain tissues upon behavior has been described by Brickner. The patient was a successful business man in whom a brain tumor necessitated the removal of large areas of both frontal lobes. The patient apparently showed little loss of intellectual ability as a result of this drastic operation. He was able to remember old experiences and to form new associations; he could, for example, memorize poems and solve mathematical problems. The patient did, however, show certain marked changes in emotional control and in personality. He was less inhibited and often silly in speech and action; also he became boastful and showed less regard for his friends and his family. In this man it appeared that the intellectual functions of the removed frontal areas were carried on by other parts of the brain. But there was a general loss in emotional control and less regard for social inhibitions.

By way of summary, it should be made clear that the concepts of equipotentiality and mass action do not mean that there is no localization of function in the brain. Certain brain areas have been definitely marked off and their functions are well known. The somesthetic and proprioceptive centers, for instance, are found in the parietal lobes and injury to these regions leads to loss of sensation in specific parts of the body. The motor areas are located in the frontal lobes and injuries here lead to loss of voluntary movements. Brain areas concerned with vision, audition, taste, and smell are well established, and studies of aphasia have shown articulate speech to be in some way connected with structures in the parietal and temporal lobes. But no center for intelligent behavior or for learning has been localized, and it is in just such complex activities that substitution and transfer of function take place. Associations and learned acts must depend upon much more extensive brain patterns than was formerly thought probable. This may explain why a single operation does not greatly disturb a given pattern, but the greater the injury the greater the likelihood of disturbance. In restitution of function it seems reasonable to surmise that associations have been reorganized following a break in the anatomical pattern.

It may seem as though we are returning to an older and less analytic view to say that the brain "functions as a whole," but the work of Franz and Lashley make this view seem more probable than it once appeared to be. The brain, to be sure, is not a sponge, the parts of which are totally undifferentiated. Localization is an established fact. But our concepts of cortical areas must apparently be modified to allow for wide flexibility, and extended to permit of much functional identity of one part

with another. The clinical as well as the experimental evidence indicates that equipotentiality found for the rat's brain holds within broad limits for the brains of cats, monkeys, and men.

7. ELECTROENCEPHALOGRAPHY OR THE STUDY OF "BRAIN WAVES"

We have seen that while many functions have been localized within certain cortical areas, the brain itself possesses great adaptability and organizing power. Further evidence that the brain cannot be thought of as an intricate telephone exchange inserted between incoming and outgoing lines, merely waiting to be activated by visual, auditory, tactual, and other stimuli, is provided by studies of "brain waves." The recording of brain waves is known technically as electroencephalography. The waves themselves are a result of spontaneous activity—much like pulsations or "beats"—within the brain itself. They may be obtained in the following way. First, a pair of electrodes is fastened to the head, one on the forehead, the other on the back of the skull (other locations will also work); and these are connected with a kind of radio set containing amplifying vacuum tubes. The continuous electrical pulsations then picked up are recorded by writing points upon a moving strip of paper as "waves" or as up-and-down fluctuations.

These continuous electrical phenomena are classified as alpha, beta, and delta waves. The alpha rhythms come at the rate of about 10 per second and are fairly regular in form. Beta waves are smaller in amplitude and faster than alphas; their average frequency is around 25 per second. Delta waves—much larger than alphas ordinarily and fairly regular—are quite slow, about 4 to 6 per second. Alpha waves are known to originate in the cerebral cortex itself while deltas arise in the underlying subcortical structures (the hypothalamus). Alpha and delta waves have been more carefully investigated than have the beta waves. See Figure 9.

To the psychologist the question of greatest interest is the correlation of brain wave changes with psychological activities. Visual attention to any situation, even shutting and then opening the eyes, is known to depress the alpha rhythms; holding the eyes closed reinstates alpha activity again. The sound of a buzzer; working on an arithmetic problem; reading a book; carrying out instructions requiring choice, cause the alpha waves to decrease markedly in amplitude and often to disappear altogether. On the other hand, an unexpected question, a ready signal, completion of a task, apparently any sudden change in a person's "mental set," is accompanied by a sudden burst of alpha waves. In general, active attention tends to depress or thin out the alpha waves,

while relaxation appears to produce renewed alpha activity. It seems a reasonable guess that the fall in alpha with increase in attention results from a sudden switching of energy from the brain centers and its redirection into muscular and sensory channels.

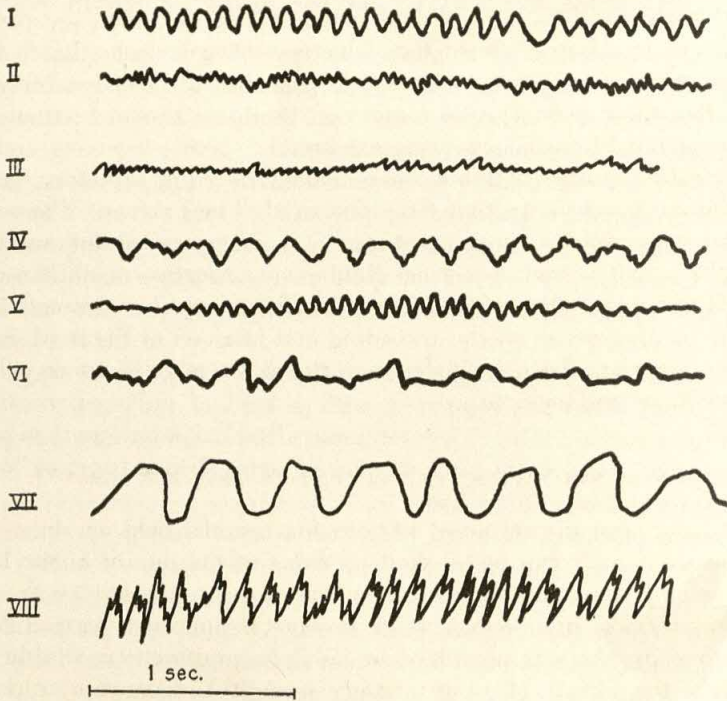


FIG. 9

GRAPHIC RECORDS OF THE EIGHT TYPES OF CEREBRAL ACTION—POTENTIAL WAVES
 I, alpha; II, beta; III, gamma; IV, saw-tooth; V, spindles; VI, large "random waves";
 VII, delta; VIII, spikes. As the length of the line labeled "1 sec." indicates, each of the
 above records covers a period of 3 secs. (From Crafts, Schneirla, Robinson, and Gil-
 bert, *Recent Experiments in Psychology*, 2nd ed. Copyright, 1950. McGraw-Hill Book
 Company, Inc.)

Decrease in alpha activity is usually accompanied by increase in beta and delta waves. This may be a balancing effect, owing to the fact that these waves and alpha have different origins within the brain; it may arise from the fact that different regions of the cortex have characteristic and different wave activity; and it may be an artifact due to the method of recording.

Other findings about brain waves are of value to the psychologist. In sleep the alpha rhythm varies markedly, and five levels corresponding to stages of sleep have been noted. Drugs, narcotics, changes in tempera-

ture and in oxygen supply also lead to characteristic and recognizable changes in brain wave activity. It has proved possible to locate brain tumors by detecting changes in the waves coming from different parts of the brain. Epilepsy and certain of the insanities are known to give distinctive brain wave patterns which will enable the physician eventually to understand these conditions better, to diagnose their severity and predict their progress. Since the delta waves come from the thalamic regions which are known to be active in emotional behavior, the possibility that these may serve as indices of emotional stress or pathological emotional conditions is now being explored.

Interesting studies of brain wave characteristics in hereditary feeble-mindedness have shown that the alpha rhythm becomes more and more normal with increase in mental age. This suggests that the variations in alpha activity found in normal children is a function of differences in mental age rather than of differences in chronological or life age. Work is now in progress upon the recording and analysis of brain waves. As these phenomena become better understood, their value to psychology and medicine will increase.

8. A SYSTEMATIC EXPERIMENT IN PSYCHOSURGERY

We have seen in preceding sections that normal behavior in man as well as in animals can be affected by experimental injury to the brain. Is it possible that abnormal behavior can also be affected (improved) by some form of brain surgery? To answer this question, the Columbia-Greystone project was organized in 1947. Its purpose was to study intensively the effects of brain surgery upon the subsequent physical, psychological and emotional condition of mental patients. Previous to this work a variety of operations performed upon the brains of the insane had led to a greater or lesser degree of improvement. Some of the patients recovered completely, others showed little or no improvement, and a few were even worse. There was hope, however, that a systematic study of different brain areas with new and better techniques might lead to a higher per cent of cures. In any event, psychosurgery seemed promising enough to warrant a thorough test of its efficacy. An interesting aspect of the Columbia-Greystone project is the fact that it was a collaborative venture. Anatomists, neurologists, physicians, psychologists and psychiatrists all contributed according to their specialties.

Subjects in the experiment were 48 psychotic patients, 30 men and 18 women. All had been in mental hospitals for at least two years and had shown no improvement—that is, no response to any form of therapy. Patients were arranged in pairs, the two members of each pair being

matched as nearly as possible for medical, psychological, and socio-educational factors. One member of each pair underwent the brain operation and the changes in his condition were checked against his unoperated (control) match. In all of the operations some part of the frontal lobe was removed, the part excised being systematically varied from patient to patient. The surgical removal of specific areas of the brain is called *topectomy*.

Extensive physical, medical, and psychological examinations were made of all 48 patients before and after the operation. Four months later the status of each patient was reviewed by hospital psychiatrists unaware of which patients were experimentals and which were controls. Four of the controls (unoperated) and 20 of the experimentals (operated) were sufficiently improved to be recommended for parole—certainly a promising result.

A detailed analysis of the effects of psychosurgery leads to the following conclusions:

- 1) There were no marked improvements in medical, physiological or neurological condition. Nor were there significant changes in the electroencephalograms, in vision, audition, or other sensory functions. Several patients had convulsive seizures from which they later recovered; and a few had temporary and transient paralyses. Although the speech area (Broca's area) was removed in two patients, no aphasia or speech disturbance was discernible.

- 2) Psychosurgery produced no measurable improvement in general intelligence, in verbal association or in tests of judgment and learning. Rorschach reports were not significantly altered. In fact, none of the mental tests showed significant differences as between operated and unoperated patients.

- 3) Most of the patients showed a transient loss in the Porteus Maze test (a measure of planning capacity, see p. 137) which was recovered in approximately four months.

- 4) By all odds the most striking and significant finding was a marked loss in anxiety as a result of the operation. This was revealed by various inventories (pp. 186–189) as well as by improvement in social relations and decrease in physical complaints.

- 5) The psychological findings as to loss of anxiety were corroborated by the psychiatric interviews. These indicated that the operated patients had achieved an improved attitude toward the world and toward themselves. Such improvement seemed similar in many ways to that change in feeling tone exhibited by a person who has recovered from a severe illness.

It is quite apparent from this study that psychosurgery is not as yet a

panacea for all mental ills. At the same time it does in general lead to improvement in a patient's condition, and often brings about drastic improvement when all other therapies have failed.

BIBLIOGRAPHY

1. Brickner, R. M. The intellectual functions of the frontal lobes. New York: The Macmillan Company, 1936.
2. Brown, C. W., & Ghiselli, E. E. Subcortical mechanisms in learning. II. The maze. *J. comp. Psychol.*, 1938, 26, 27-44.
3. Columbia-Greystone Associates. Selective partial ablation of the frontal cortex [Mettler, F. A. (Ed.)]. New York: Paul B. Hoeber, Inc., 1949. Section III, especially.
4. Franz, S. I. On the functions of the cerebrum: the frontal lobes. *Arch. Psychol.*, New York: 1907, 1, No. 2.
5. Franz, S. I. Nervous and mental re-education. New York: The Macmillan Company, 1923.
6. Harlow, J. M. Passage of an iron rod through the head. *Boston Med. & Surg. J.*, 1848, 39, 389-393.
7. Hoagland, H., Cameron, D. E., & Rubin, M. A. Emotions in man as tested by the delta index of the electroencephalogram. *J. gen. Psychol.*, 1938, 19, 227-245.
8. Jacobsen, C. F. Functions of frontal association area in primates. *Arch. Neurol. Psychiat.*, 1935, 33, 558-569.
9. Jacobsen, C. F., & Taylor, F. V. Behavioral changes associated with lesions of the frontal lobes in monkeys. *Psychol. Bull.*, 1937, 34, 767.
10. Jasper, H. H. Electrical signs of cortical activity. *Psychol. Bull.*, 1937, 34, 411-481.
11. Jasper, H. H. Electroencephalography. Chapter 14 in Penfield, W., & Erickson, T. C. *Epilepsy and cerebral localization*. Springfield, Illinois: C. C. Thomas, 1941, 380-454.
12. Klüver, H. Behavior mechanisms in monkeys. Chicago: University of Chicago Press, 1933.
13. Kreezer, G. The electro-encephalogram and its use in psychology. *Amer. J. Psychol.*, 1938, 51, 737-759.
14. Lashley, K. S. Studies of cerebral function in learning. *Psychobiol.*, 1920, 2, 55-135.
15. Lashley, K. S. Nervous mechanisms in learning. In Murchison, C. (Ed.), *Handbook of general experimental psychology*. Worcester, Massachusetts: Clark University Press, 1934, chapter 10.
16. Lashley, K. S. Factors limiting recovery after central nervous lesions. *J. nerv. ment. Dis.*, 1938, 88, 733-755.
17. Pike, F. H., & Chappell, M. N. On the recovery following lesions in the cerebral cortex. *Science*, 1930, 71, p. 76.
18. Williams, A. C. Some psychological correlates of the electroencephalogram. *Arch. Psychol.*, New York, 1939, 34, No. 240.

3

THORNDIKE'S LAWS OF LEARNING

1. THORNDIKE'S EXPERIMENTS IN ANIMAL LEARNING

EDWARD L. THORNDIKE'S studies in animal psychology marked the beginning of the modern experimental attack upon problems of learning and habit formation. Thorndike began his work at Harvard, where as an undergraduate he did several experiments with chicks in the cellar of William James's home. Later, he came to Columbia to work in Cattell's laboratory; and in 1898 received the doctor's degree in psychology, his dissertation embodying a series of studies under the title *Animal Intelligence: An Experimental Study of the Associative Processes in Animals*. Although he retained his interest in animal psychology, for the last forty years of his life Thorndike worked mainly in the field of educational psychology, where he carried on a number of important studies.

Prior to Thorndike much of the work with animals of such investigators as Lloyd Morgan, Lubbock, and Romanes, can rightly be called experimental even though performed without adequate laboratory control or special apparatus. For the most part, however, writers on animal psychology contented themselves with anecdotes and with uncontrolled although highly interesting observations of the doings of some pet dog or horse. Many of these accounts are impressionistic and highly colored, with little or no emphasis upon the slowness shown by most animals in learning. To be sure, animals were rarely credited with the ability to "reason" or "think"—this usually being considered exclusively a human accomplishment—but they were regarded as equal to many men, and superior to some, in their ability to form associations and build up habits. That they learned many things through imitation and the use of "ideas" seems not to have been questioned. As we shall see later, Thorndike's studies furnished experimental evidence which called into question these somewhat naïve views.

Although carried out with animal subjects, Thorndike's experiments have broad and important implications for human learning. Certain definite principles, for instance, have emerged from his studies which

characterize quite accurately the way in which children and adults form new associations and learn new things. Thorndike's two fundamental laws of learning, the *law of exercise* and the *law of effect*, in spite of criticism and emendation, are still important in modern pedagogical theory. And it seems highly probable that they will long be regarded as among the most valuable empirical generalizations with which psychologists must deal.

2. EXPERIMENTS WITH FISH AND CHICKS

Thorndike's experiments dealt with the learning ability of fishes, chickens, cats, dogs, and monkeys. The fishes, a common variety resembling the minnow, were kept in an aquarium measuring four feet by two feet, containing about nine inches of water. These fishes shun sunlight, and one end of the aquarium was covered over to protect them from light; here all food was given, and here the fishes remained most of the time. An experiment consisted in gently forcing a fish from the shady to the sunny end of the tank, by gradually moving across the tank a glass slide placed behind the fish to cut off its retreat. A second glass slide containing a small opening was then placed between the fish and the shady end of the tank. The object was to see whether the fish could find the opening and escape. At the beginning, the fish, motivated by its aversion to light, swam up and down the length of the slide, bumping against it here and there, and looking for a place to get through. Eventually it came to the opening and swam through to the shady section. Upon being replaced in the tank again and again, a fish soon indicated clearly that it had profited by previous experiences, that is to say, it swam more directly to the opening, with less loss of time and fewer random movements. This experiment was repeated with a number of fishes, and with slides containing openings at different points. Always the result was the same—in every case the fish ultimately learned the trick of finding and getting through the opening in the slide. From this experiment it is clear that learning, in the sense of simple connection-forming, can be experimentally demonstrated in vertebrates rather low in the evolutionary scale.

In his experiments with chicks, Thorndike used a number of pens or mazes, one of which is illustrated in Figure 10. When a chick is placed in section A of the maze, there are four possible exits as shown in the drawing. If the opening on the extreme right is followed by the chick to the second turn, it leads out of the maze and into an enclosure in which are other chicks and food. The other three exits are blind alleys. The problem is to see how long it will take a chick to select the pathway which will carry it out of the maze. The behavior of a chick when first placed

in the pen resembles closely that of a fish when trying to get out of the sunlight. Taken away from the other chicks and from food, and dropped

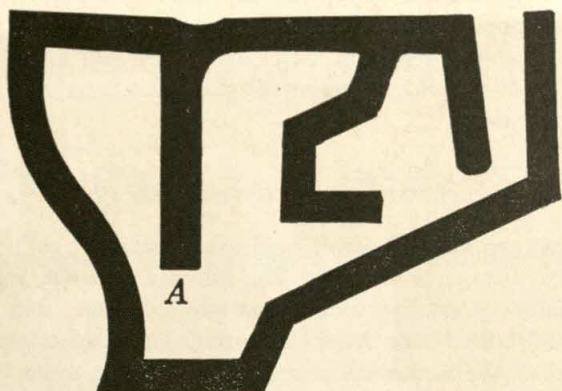


FIG. 10

PEN USED BY THORNDIKE IN STUDYING THE PROCESS OF LEARNING IN CHICKS

into the pen at A, the chick runs back and forth, in and out of the blind alleys, peeps loudly, tries to jump out of the pen and to squeeze through any available opening. At length "by accident" it picks the right exit and gets out. Put back again and again, for the first few trials the chick's behavior is much as before; soon, however, it begins to eliminate useless movements, such as repeatedly entering the blind alleys, until finally it runs directly to the right exit. A good picture of how a chick learns the trick of escaping from such a pen as the one described is shown in Figure 11. In this figure the separate trials are laid off at equal distances along the horizontal axis, and the time required to escape is laid off on

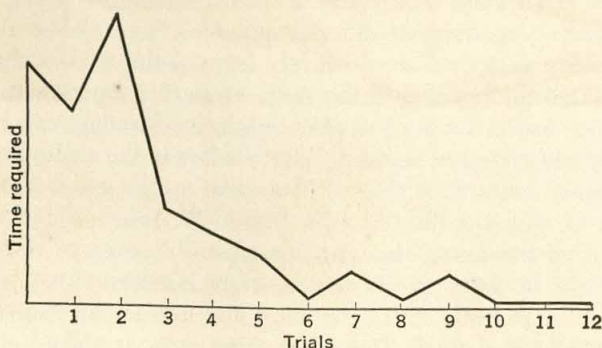
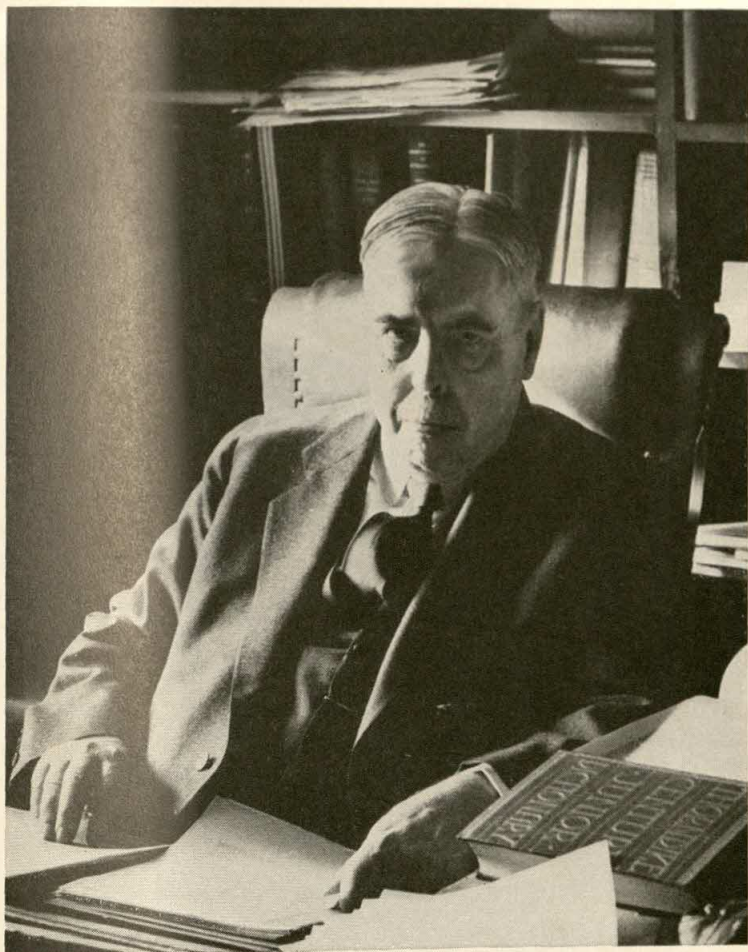


FIG. 11

LEARNING CURVE OF A CHICK

This curve shows the time required in successive trials to escape from the pen shown in Fig. 10. (From Thorndike, *Animal Intelligence*, 1911.)



EDWARD L. THORNDIKE

1874-1949

the vertical axis. As shown by the drop in the curve, after five or six trials the chick runs at once to the proper exit.

3. EXPERIMENTS WITH CATS, DOGS, AND MONKEYS

Thorndike's experiments with cats are his best known and most often quoted studies of animal learning. The task set for the cats was to escape from various "puzzle boxes," the general construction of which may be seen from Figure 12. These boxes were so designed that escape from

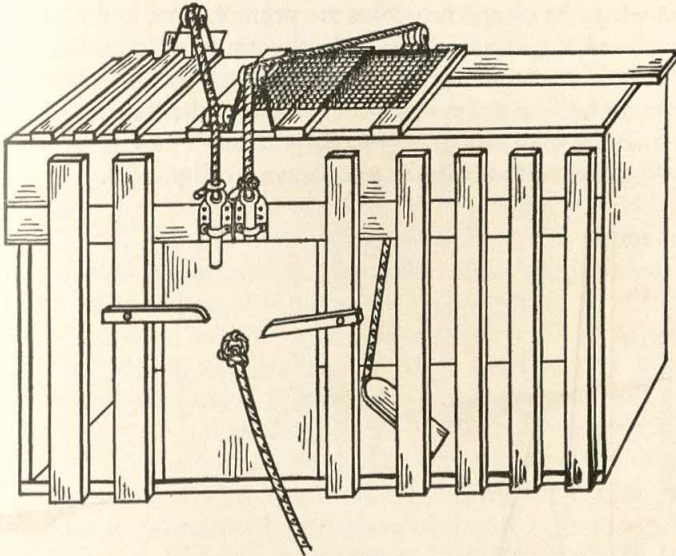


FIG. 12

PUZZLE BOX USED BY THORNDIKE IN LEARNING EXPERIMENTS WITH CATS

them could be effected in a variety of ways: e.g., by turning a button, pulling a string, depressing a lever, or pulling a wire loop. Each of these escape mechanisms, when operated by the animal, released a door which was at once pulled open automatically by a weight attached to it. Only one escape device was employed with the simpler boxes; with the harder ones, two or more separate acts, such as pulling a string and depressing a bar, were required to open the door. An experiment consisted in placing a hungry cat in the box with a small piece of fish or meat lying just outside. This situation usually resulted in immediate activity on the part of the cat. It tried to squeeze through the slats of the box; clawed at the bars of wire; thrust its paw through any opening large enough; worked vigorously at anything loose or movable, and in general gave a

perfect picture of impulsive hit-or-miss effort to escape. In time the animal nearly always succeeded in operating the escape device (hitting the wire noose or turning the button) by accident, and getting out. It was then allowed to eat a bit of fish or meat, and immediately returned to the box for a second trial. During this next trial, and for several trials thereafter, the cat's plan of attack remained much the same as before; but in succeeding trials its activity became more and more restricted to the button or other escape device, and useless clawings and scratchings were gradually eliminated. Finally the animal, on being placed in the box, went almost at once to the door, worked the mechanism, and escaped.

A cat may take twenty or more trials and require an hour or so to reach the point where its escape-responses are prompt, sure, and accurate. During this process, improvement is nearly always quite irregular, the escape time seesawing up and down until the act is well learned, whereupon the time becomes fairly regular. A learning curve of the time required in successive trials by one young cat to escape from a box kept closed by a single bolt attached to a wire loop is shown in Figure 13.

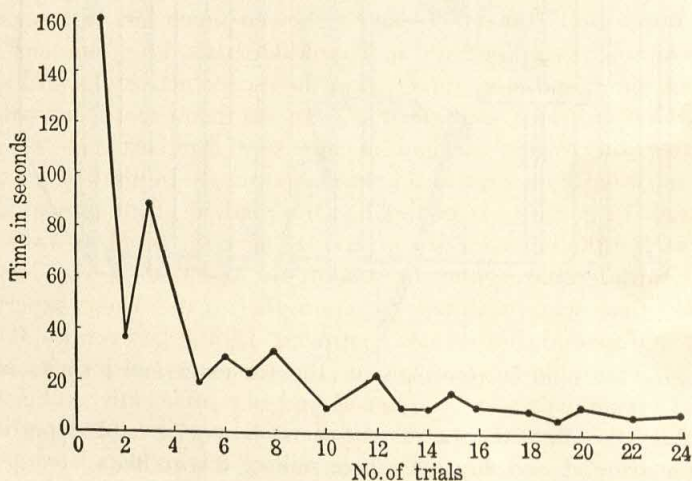


FIG. 13

LEARNING CURVE OF A CAT

The time taken by the cat to escape in successive trials is represented by the height of the curve; the successive trials are laid off along the horizontal baseline. (From Thorndike.)

In the figure the twenty-four separate trials are represented by equal distances laid off along the horizontal axis; and the time to escape in seconds is laid off on the vertical axis. Note the jagged appearance of the learning curve in the earlier trials as the time is first longer and

then shorter; and the smoother appearance of the performance line as the cat learns the trick. This graph is fairly typical of animal performance in situations like the one described, and is a good instance of what is generally called "trial-and-error" or "trial-and-success" learning. Such learning begins as a varied hit-or-miss process and continues as such until the successful response is hit upon, as we say, "by accident." After this, elimination of the unsuccessful responses begins, together with a gradual building-in of the successful reactions. The cat whose performance is shown in Figure 13 was evidently learning in a slow and painstaking way to reduce its useless activities. At the same time the successful response of pulling the wire loop was becoming more and more firmly established.

Thorndike's account of how the cat escapes from the puzzle box has become the classical description of animal learning in situations like the one described. It has not escaped criticism by later experimenters, however, who object that cats exhibit much greater ability to think their way out of puzzle boxes than they are given credit for by Thorndike. In one careful experiment, in which Thorndike's boxes and conditions were closely duplicated, the cats' behavior showed much less excitement and "mad scramble" than appeared in Thorndike's cats. Less "random" movement was found and more direction of the escape activity toward specific objects (button, door, and the like). In another experiment with cats and boxes, the escape mechanism was kept standard and the act of escape recorded by a camera operated electrically by the escape mechanism itself. In getting out of this box the method of attack varied from cat to cat, but the behavior of a given cat (once the problem was solved) showed considerable rigidity or stereotypy. Again, there was less "randomness" than was exhibited by Thorndike's cats. These experiments amplify and amend, but do not controvert Thorndike's results. It seems likely that Thorndike overemphasized the aimlessness of his cats' activity and that his animals were frightened and not sufficiently acclimated to the puzzle box situation. Later reviewers, however, tended perhaps to stress the excited and dramatic behavior of Thorndike's younger cats rather than the quieter behavior of his older and tamer cats. On the whole, it appears that trial-and-error is still the best characterization of feline learning in situations where the problem is new, strange, and not closely related to the animal's previous experience.

In some experiments with dogs, Thorndike used puzzle boxes of the same sort as those used with cats. The dogs, like the cats, employed the method of trial-and-error in trying to get out, but on the whole their learning curves tended to be smoother. This indicated to Thorndike better observation, a more planful attack, and perhaps a somewhat higher level

of intelligence. That dogs learn more quickly than cats, and somewhat less quickly than monkeys or raccoons, seems to be the consensus of investigators who have worked with these animals. However, the inferiority of the dog to the raccoon or even the monkey is not well established, since in many of the tests the dog is at a disadvantage because his paws cannot so easily manipulate mechanical devices as can the digits of the raccoon or monkey. Moreover, we cannot be certain that different animals have the same motivation at the time they are tested.

Experiments much like those described on dogs and cats were also carried out by Thorndike upon three small South American monkeys. These monkeys were tested to see how readily they could learn to operate simple mechanical devices so as to get into or out of boxes. In addition, they were taught to manipulate a simple mechanism by means of which food could be brought into their cage. The results of these tests are interesting and quite definite. In almost every instance the monkeys quickly learned the trick—often at an astonishing rate, far surpassing in skill and speed both the dog and the cat. The superiority of the monkey in tests of this sort is probably to be expected, because as Thorndike points out, they have better vision than dogs or cats; and this advantage, plus the fact that they have fingers, would enable them to manipulate simple contrivances more easily than other animals. Besides these natural assets, monkeys are more active and more curious than other animals, being quicker to play with and manipulate movable objects.

4. DO ANIMALS THINK OR REASON IN SOLVING PROBLEMS?

In spite of the superiority of monkeys over other animals both in speed and permanence of learning, they exhibited, says Thorndike, no real understanding of, nor "insight" into, the problem to be solved. This appeared to be even more nearly true of other animals. In fact, throughout his experiments Thorndike did not find any clear evidence that an animal ever "thinks its way through a problem," i.e., observes relations clearly, uses "ideas," makes inferences and comparisons. This conclusion came out most clearly in a series of experiments involving choice and discrimination. The monkeys had formed the habit of coming down to the bottom of the cage from their accustomed place near the top whenever the experimenter appeared with food. This habit was used to test the monkey's power of discrimination, in the following way. Whenever the experimenter picked up food with his left hand, the monkey was fed; but when he picked up food with his right hand, the monkey was never fed. At another time food was given when taken from a brown pasteboard box, and withheld when taken from a white crockery cover; or food was

given at one visual signal (a large letter or geometrical figure), and withheld at another. The question was whether, after having learned to discriminate between food and the non-food signals, the monkey would get the general idea: *one stimulus means food, the other does not*. Thorndike argued that evidence of insight or understanding into the situation might be expected to show itself if, after several discriminations of the kind described, the monkey perceived the relation almost instantly. On the other hand, if there were no clear abstraction of the principle of choice required, the monkey might be expected to take each discrimination as a new task to be solved by trial-and-error. This would, of course, show up in the large number of trials needed for each new task. The results of these tests were not entirely unambiguous, but their general implication seemed to be clear. The monkeys learned some of the discriminations very quickly, but failed to learn others. Thorndike concluded that they had not formed a clear concept of what was essentially involved in *all* of the tasks which they could use in governing their future behavior.

Later experimenters have disagreed with Thorndike's conclusions that monkeys (and other animals also) do not solve problems by reasoning and the use of "ideas." Part, at least, of this disagreement is undoubtedly a matter of definition. If by reasoning one means the use of symbols—verbal and mathematical formulations—then animals either do not reason, or it would be hard to prove that they do. If, however, by reasoning one simply means the combining into one action (for example, of escape) of habits acquired upon different occasions in the past, then animals certainly do appear to reason. Maier found that white rats trained in two separate motor habits could later combine these experiences into one integrated and successful response, namely, getting food placed behind an obstruction. To be sure, this experiment presents some difficulties in interpretation. One cannot be certain that the habits were unrelated, or that the "new" behavior was actually a synthesis of the previous separate experiences. Another attack upon the problem of reasoning in animals has been made through use of the *multiple choice method* devised by R. M. Yerkes. The general plan of the multiple choice apparatus is shown in Figure 14. The problem set the animal may best be made clear, perhaps, by an illustration. Suppose, first, that the doors to boxes 1, 2, and 3 are open, and that box 2 contains food. Next, doors to 7, 8, and 9 are open and box 8 contains food; or boxes 3, 4, 5, 6, 7 are open and box 5 contains food. The task is to learn the constant relationship within the different situations. If the animal gets the relations involved (food = middle door) it readily solves the problems. Other problems investigated by the multiple choice method are those in which the

food box is (1) at the right or left end of the series; (2) is second from the left or right; (3) is alternately at the left or the right end. Discriminations of a high order are demanded by the hard problems. Comparative studies of various species have shown the monkey and the ape to be far superior to the rat and to most birds in ability to solve problems of the sort described.

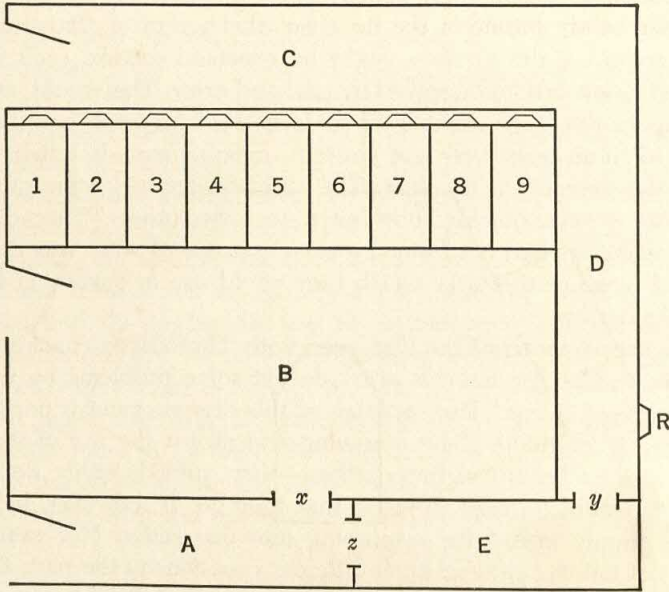


FIG. 14

PLAN OF MULTIPLE CHOICE APPARATUS

The animal is introduced at A, then released into B, the response compartment. It enters one of the reaction boxes, numbered from 1 to 9. If the choice is correct, the animal finds food in a cup at the rear of the box, or it may be released into C and D to find food at R. If the choice is incorrect, the animal may be confined for a short time in the box it enters, as punishment. From D the animal enters E to await the next trial. Sliding doors are provided at x, y, z, and at the front and rear of each reaction box. The entire apparatus is about 18 feet wide and 20 feet long. (Modified from Yerkes, p. 5.)

The literally tremendous advantage held by men over animals because of the possession of language is clearly shown in multiple choice problems. Once the appropriate discriminations are made, a man need only say "second from left" silently or aloud when the next combination of boxes is presented—trial-and-error is reduced to a minimum. The advantage of language comes out clearly also in another well-known laboratory experiment—that involving "delayed reactions." The plan of an apparatus used in delayed reaction study is shown in Figure 15. When the

animal is released from compartment *E* (which has glass at the front and on the sides), it learns, after a number of trials, that the lighted box contains food. The box selected for lighting varies, of course, in random order from trial to trial. The food in the lighted box furnishes the positive incentive for learning; a weak electric shock, which the animal gets before the unlighted doors, is the negative incentive. After an animal has learned the association (lighted box = food) it is placed in *E* and the light flashed as before in one of the three boxes. Then the light is turned off,

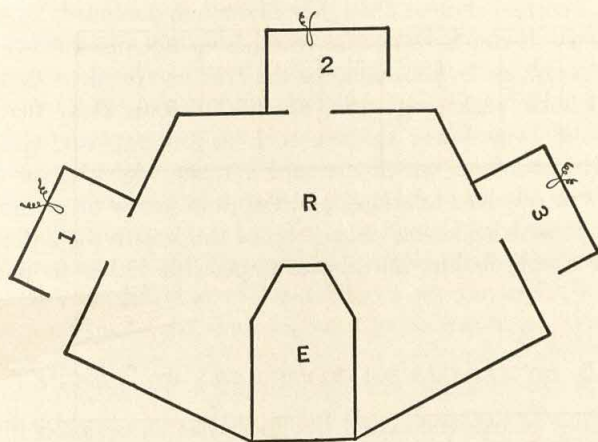


FIG. 15

PLAN OF DELAYED-RESPONSE APPARATUS

The animal is released from *E* into *R*, the response compartment, and then enters one of the reaction boxes, number 1, 2, and 3. These are lighted at the rear. (Modified from Hunter, p. 24.)

the animal meanwhile being restrained for varying intervals in *E* before being released. The purpose of the delay is to see how long the animal can be held and still find the right box. Since the reaction cannot be to the light which is no longer present, it must be to some stimulus-substitute, an "idea" or other memory cue, perhaps a muscle- or postural-set. Comparative studies of the maximum period of delay, which is followed by success, have been made of white rats, cats, dogs, raccoons, monkeys, and apes. Conflicting results are reported owing in part, no doubt, to variations in experimental setup. In several experiments, for instance, the food itself instead of the light is the signal; there is no training period in which the animal learns that light "means" food. Apes and monkeys can react successfully after longer periods of delay than can cats, dogs, or rats; but the white rat makes a surprisingly good showing

as compared with dogs and cats. The delayed reaction technique has been successfully used in similar experiments with young children.

One interesting finding in delayed reaction experiments is the importance of body and head orientation to rats and dogs. After the light has flashed, these animals tend to keep the body or head oriented (pointed) toward the box in which the light was last seen. When released, the animal then simply "follows its nose" to the food box—no memory cue is needed. If the animal's "position set" is disturbed—if, for instance, the animal while waiting is turned around or forced to move—it is often still able to make correct choices, though at first much confused.

In summary, it can be said that whether or not one believes that animals solve problems by reasoning or the free use of ideas depends upon how one defines "reasoning" and "ideas." It seems clear that apes and monkeys—and even lower species, such as the rat, cat, and dog—are able to solve problems which demand a fairly high degree of abstraction, ability to discover relations, and the persistence of memory impressions. That animals make use of symbols or think with the aid of language as men do is improbable, but almost impossible either to prove or disprove.

5. DO ANIMALS SOLVE PROBLEMS BY "INSIGHT"?

Another line of evidence cited by Thorndike as opposed to the thesis that animals think or reason is the fact that few sudden drops occurred in the learning curves of his cats, dogs, and monkeys. According to Thorndike, sudden drops show that the learner has got the point—seen the connections or relations—and henceforth may be expected to solve the problem whenever the situation presents itself. Instead of this sudden insight, what Thorndike found was a gradual sloughing-off of excess and useless movements (see Figures 11 and 13), with no clear evidence that the animal observed how it got out or made use of these observations in future trials.

Thorndike's view that animals learn almost if not entirely by hit-or-miss methods, and evince little insight into the problem at hand, has been attacked by Köhler and Koffka, who belong to the school of Gestalt psychology.¹ In his studies of learning in the chimpanzee, Köhler found many instances of quick learning which, he says, indicate that the ape suddenly grasped the relations involved in the problem. According to Thorndike, such quick learning (which presumably involves insight) is to be expected when the task is "very simple, very obvious, and very clearly defined"; but whenever the problem is at all complex, the ani-

¹ See Chapter 4.

mal's behavior, he thinks, may be fairly described as "stupid." In Koffka's view, on the contrary, the animal exhibits insight or intelligence whenever it is possible for it to grasp the problem; stupid errors, he says, occur when the task seems simple to us, but is almost surely not simple to the cat or chick. Koffka argues further that Thorndike's puzzle boxes set before the animals tasks so difficult (for the animal) that trial-and-error was the only possible method of attack. In spite of the difficulty involved in these tasks, Koffka points out instances in Thorndike's own data where sudden vertical drops occur in the learning curve and these, he suggests, indicate insight into the problem.

It is probable that Thorndike overemphasized the stupidity of the animal attack upon a problem, partly, no doubt, because of the *kind* of problems which he set for his animals. And it is equally probable that Köhler found insight in his apes because of the kind of problem which he set. Köhler's chimpanzees were faced by tasks very different from those required of Thorndike's monkeys and cats, and for this reason the two sets of results are not directly comparable. For one thing, the chimpanzees were given much freedom, while Thorndike's animals were nearly always confined. Examples of the kind of tasks Köhler set the chimpanzees were (1) securing a banana suspended from the ceiling of a cage by piling up boxes one on another; (2) reaching for and pulling into the cage a banana placed outside by ingeniously hooking together two sticks. Many other tasks involved ropes, sticks, and the piling-up of boxes.² In such situations sudden learning might very well take place, when the animals are intelligent to begin with (as Köhler's chimpanzees undoubtedly were) and when the task is not too different from the kind of things which the animals habitually do.

Another point to be considered is the fact that *all* of the learning is not shown in the fluctuations of the learning curve. A sudden drop in the curve may be preceded by a long trial-and-error process, which is not represented in the curve, as when a man suddenly "sees the point" of a problem because of long and varied experience with similar problems. The chimpanzee who suddenly does a trick which he could not do before has not necessarily fathomed it in a sudden burst of comprehension. Many tentative trials not shown in the learning curve and not seen by the experimenter may have preceded the solution. The importance of past experience has been demonstrated in an experiment in which chimpanzees were unable to solve a problem, using sticks as reaching tools, unless they had previously played with sticks, joined them together, and used them in various ways. Maier, who has studied reasoning in animals and in man, suggests that the solution of a problem demands (1) *selec-*

² See Chapter 4.

tion as a result of past experiences with elements of the problem, and also (2) *direction*—an overall idea or plan of attack. It is the latter which animals seem to lack most often.

6. DO ANIMALS USE IMITATION IN SOLVING PROBLEMS?

Before considering the experimental evidence for and against imitation in animals, it will be well to have clearly in mind just what is meant by the term "imitation." Suppose dog A barks and sets out after a passing cat, and is followed immediately by dog B who also barks and chases the cat. Is B imitating A? The proud owner of A will probably say yes, although it is more likely that B is simply responding to the same stimulus (moving cat) as is A. Imitation is present when a person or animal performs an act which is stimulated or instigated by observing similar behavior in another person or animal. To be genuine imitation must satisfy the following criteria: (1) the act to be imitated must be unfamiliar, and not already part of the imitator's repertory; (2) the imitative behavior must follow immediately upon observation without evidence of trial-and-error behavior; (3) the imitative behavior must be a recognizable copy of the observed behavior. When learning occurs promptly under these conditions, the process may be fairly described as imitation.

Thorndike's statement that his animals gave little evidence of learning through imitation provokes astonishment and protest from many owners of pets. Yet his results point directly to this conclusion. When a cat, for instance, which had not learned to get out of a puzzle box was allowed to observe another trained cat, it made no difference whatever in the first cat's behavior. Nor did it make any difference when an untrained cat was placed in the box with a trained cat and allowed to escape when the second cat opened the door of the cage. The untrained cat still used the old hit-or-miss method as before; there was not the slightest evidence that learning had been speeded up through observation. This same result was obtained in similar experiments with dogs and chicks. Even monkeys who had failed in their efforts to operate some simple mechanism were unable to do it after having seen a monkey or the experimenter do the trick many times. In one case a monkey was shown fifteen consecutive times how to open a box which contained food (the door was held shut by a simple lever). At the end of this training period, the monkey's efforts did not differ essentially from its previous attempts before tuition. In summing up his observations on imitation, Thorndike wrote: "Nothing in my experience with these animals, then, favors the hypothesis that they have any general ability to learn to do things from seeing others do them."

One may well inquire—if these observations are accepted—why monkeys are so generally believed to be accomplished imitators. Thorndike's answer is that monkeys, being active, curious, and possessed of a repertoire of movements much like ours, will inevitably do many things which seem almost human. He writes, "If you put two toothpicks on a dish, take one and put it in your mouth, a monkey will do the same, not because he profits by your example, but because he instinctively puts nearly all small objects in his mouth. Because of their general activity, their instinctive impulses to grab, drop, bite, rub, carry, move about, turn over, etc., any novel object within their reach, their constant movement and assumption of all sorts of postures, monkeys perform many acts like our own and simulate imitation to a far greater extent than other mammals."

More recent studies of imitation in monkeys have questioned the sweeping character of Thorndike's findings. In one careful study of imitation in the monkey, two cages were placed side by side. In one cage was a monkey which had been trained to do various tasks: pulling a chain hanging down in front of the panel; opening a door in the panel by turning a knob; pushing down two latches and opening the door, and the like. In the other cage was an untrained monkey. Care was taken to give the untrained animal a clear view of the trained, as these investigators believed that Thorndike's animals were not given sufficient opportunity to observe the act to be copied. Now if the untrained monkey copied the trained one successfully within the one minute interval allowed—without fumbling or trial-and-error—imitation was judged to have taken place. In 70 per cent of the tasks set, the act was successfully performed—often within a few seconds. Contrary to Thorndike's results, these experiments indicate that—under carefully controlled conditions—the monkey is capable of a high degree of imitative behavior. But a really favorable opportunity to imitate must be provided.

7. FIVE PRINCIPLES OF LEARNING

Thorndike has laid down five principles which serve to characterize animal learning. He regarded these principles as somewhat less important than and subordinate to the laws of exercise and effect (p. 56). They are as follows:

1. *Multiple response*, or varied reaction to the situation. Before the appropriate response can be rewarded, it must occur. Ordinarily, the animal brings to bear all of the instinctive and learned responses of which it is capable when faced by a new problem not fully understood. Multiple response usually implies trial-and-error learning.

2. *Set, Attitude, or Disposition*. Set, or attitude, is the internal "drive"

or condition which predisposes the animal to one particular kind of behavior rather than some other kind. A hungry cat, for example, will try to escape from a puzzle box if food is outside, while a well-fed cat is content to remain quiet, unless much frightened or disturbed.

3. *Partial or Piecemeal Activity*. As learning proceeds, the animal gradually eliminates useless and unsuccessful movements and confines its activities to the more promising objects or details or to those which have hitherto proved to be of value. From tentative responses, effort becomes more restricted and specific. This is essentially the concept of discrimination discussed on page 7 in relation to conditioning experiments.

4. *Assimilation or analogy* means that an animal when put into a new and strange situation will draw upon those activities which have proved successful in like or somewhat like situations. Response by analogy involves the same idea as stimulus generalization (p. 8) or stimulus equivalence.

5. *Associative Shifting*. This principle is the same as that involved in instrumental conditioning (p. 14) and it might also be taken to cover classical (Pavlovian) conditioning. The essential feature is a shift of a (reinforced) response from an effective to a previously ineffective stimulus through generalization. Ordinary animal tricks by the score furnish illustrations of this principle. The cat taught first to come when a saucer of milk is held in the hand will, after a few times, respond to the sight of the saucer or even to the person who does the feeding. In like manner a dog or monkey can be taught to sit up and beg when a command only is given. It is said that bears can be taught to "dance to music" (i.e., lift their feet) by being placed upon hot grills while music is played. Shifting of response from one stimulus to another, or from the total situation to some part of it, enormously expands the animal's range of responses. This increase in range of response is of fundamental importance in human as well as in animal learning (pp. 14-17).

One of the most striking instances of seemingly intelligent learning in an animal—explainable in terms of associative shifting—is that of the famous German horse, Clever Hans. Hans, a five-year-old horse, was trained by his owner Herr von Osten to answer questions and solve problems. A letter system was constructed on a numbered chart to enable Hans to give his replies; if the letter *m*, for example, occurred in the fourth vertical column and the second horizontal row of the chart, four taps with the left hoof and two with the right meant *m*. In solving numerical problems, taps with the right hoof meant units and taps with the left hoof meant tens. Thus 53 was represented by five taps with the left hoof and three with the right. By means of these signals, Hans was able to demonstrate his ability to solve problems and answer questions ordinarily re-

quiring considerable intelligence. The wide interest aroused in Hans' performance led finally to his being examined at successive times by two commissions. A psychologist on the second commission named Pfungst apparently solved the mystery of Hans' answers. He showed that the horse responded to involuntary "cues" or signals given by his trainer, such as slight (approving) movements of the head, changes in facial expression, and the like. Thus if the answer to a problem were 68, Hans would count rapidly with his left hoof until some small (and almost certainly unintentional) signal from his trainer warned him to stop and begin with the other hoof. The trainer, therefore, was himself the substitute stimulus to which the clever horse responded.

Hans was later acquired by another owner and became one of the group of famous Elberfeld horses. One of these horses, an Arabian named Mohammed, was reputed to be able to extract square and cube roots besides performing the simpler arithmetic operations of addition, subtraction, multiplication, and division. No satisfactory proof seems to have been brought forward to show that these horses were responding to cues, but such seems undoubtedly to have been the case. For one thing, these horses learned too quickly to signify any real understanding of the problem; they took no longer to solve hard problems than easy ones; often they would begin tapping without even glancing at the problem set up before them. Again, the kinds of mistakes made were not those of an intelligent calculator; common errors were reversals of figures such as 36 for 63, or errors of one unit such as 26 for 25. These errors might easily happen if the horse temporarily confused the left and right hoof or failed to stop tapping quite soon enough. All of the evidence, then, is emphatically against the opinion that performances of clever animals really indicate mathematical genius or intelligence of a high order. Such "stunts" do show that animals are extremely clever in picking up very slight cues and signals not readily observed by onlookers, who, of course, are not usually watching out for them.

All of the principles outlined above apply with little modification to human learning. When placed in an unfamiliar situation or faced by a complex and little-understood problem, the behavior of a man will differ very little from that of a dog, cat, or monkey. His activity at first is largely trial-and-error, as he draws upon his repertory of past experiences for a possible solution to his immediate difficulty. Multiple response passes over into piecemeal activity when certain features of the problem offer greater hope of success than others. Associative shifting is common in human learning. Many illustrations come to mind: a young man has an unreasonable distaste for orange juice because at some time in the past orange juice was used as a "disguise" for castor oil; a child is fearful of the dark after having

heard fantastic ghost stories; a college student cannot listen to hymns because of their association with his mother's funeral. Teaching a child that *C A T* "means" a small furry animal, and that notes on a chart "stand for" keys on a piano, are examples of associative shifting (see also p. 54).

A good illustration of human learning in situations which are roughly analogous to Thorndike's puzzle boxes is found in some experiments with difficult Chinese ring puzzles. Learning curves for these puzzles look very much like those of Thorndike's cats, dogs, and chicks; but instead of the irregular rise and fall observed in animal learning curves, sudden drops are frequent. These, according to the subjects' reports, occurred when the mechanical features or principles of the puzzle were suddenly comprehended.

There is a distinct gap between the things which men and animals *can* learn and the *rates* at which they learn them. Very soon in the course of human learning, often at the very beginning, men begin to use verbal formulations. Alternatives are accepted or rejected symbolically, rather than by actual trial-and-error; principles are perceived and applied; comprehension into the problem is gained. The ability to formulate general principles, evolve concepts, and educe abstract relations would seem to be strictly a human accomplishment.

8. LAWS OF LEARNING: THE LAW OF EXERCISE

Out of his experiments with animals Thorndike deduced his two fundamental laws of learning, *the law of exercise* and *the law of effect*. These formulations developed logically from the traditional laws of association (pp. 2-3). The law of exercise, often called the law of habit formation, has two parts, one of which is the complement of the other. The first part, *the law of use*, may be stated as follows: When a given situation is frequently followed by a certain response or group of responses, the bond or linkage between stimulus and response becomes stronger through the exercise so obtained. Operation of the law of use is clearly seen in the learning curves of Figures 11 and 13. When once the chick or the cat has learned the knack of escaping from a box or pen, its movements become faster and smoother as they are more and more often repeated.

The opposite of the law of use is *the law of disuse*: When a given situation is rarely followed by a certain response, the association between the stimulus and this response is weakened, its degree of weakness depending upon the amount of neglect. For instance, the tendency to enter a blind alleyway—strong at the beginning of the learning period—becomes progressively weaker through the cumulative effect of disuse.

Numerous examples of the law of use in human learning at once come

to mind. Learning to skate, to drive an automobile, or to use a typewriter consists in getting the right movements in their proper sequence *first*, and then rehearsing them until they run off rapidly, smoothly, and accurately. This is true also of school learning. In order to speak French, memorize a poem, or learn a set of mathematical formulas, it is necessary to repeat (i.e., exercise) the elements of the task in their right order regularly and often.

Thorndike's formulation of the law of exercise has been criticized by many psychologists as being too mechanical, and as not giving adequate consideration to other factors in the situation—motives, interests, special training, and the like. Studies of *incidental memory*, that is, memory for things and events not specifically attended to, have shown how little one may really know about many things constantly seen around him. A man looks at his watch many times a day, but cannot tell (when asked) whether the numerals on the watch face are Roman or Arabic; many people cannot report at all accurately the color or size of a paper dollar, or the number or color of the license plates on their cars; others cannot even describe their cars with fair accuracy. The school room offers much evidence, too, of the ineffectiveness of mere repetition. The story is often told of the boy who had formed the bad habit of saying "I have went." In order to correct this habit, he was required by his teacher to remain after school and write "I have gone" one hundred times. Having completed his task and finding the teacher gone, he left the following note: "I have written 'I have gone' one hundred times and since you are not here, I have went home." This boy evidently did not see any connection between *writing* "I have gone" as a chore and *saying* "I have gone"; and exercise without understanding had been of no value. Students do not acquire useful skills by sheer mechanical repetition. The examples given above—and many more might be cited—demonstrate clearly that mere frequency of the stimulus-response connection is not in and of itself sufficient to produce learning.

The law of exercise has been subjected by Thorndike to a rigorous series of experimental tests and as a consequence has been considerably modified. "Exercise" is now best regarded as a principle which describes how *skill* is acquired under certain conditions, rather than as a "cause" of learning. Exercise, trial-and-error, multiple response are useful insofar as they give the law of effect a chance to operate. Conditions for efficient learning, however, much include not only repetition on the part of the learner, but interest, satisfaction, and belief in the appropriateness and value of the task as well.

Under the law of exercise belong several sub-laws or corollaries which readily follow from it and are often useful descriptions of the conditions necessary for effective learning. These are the principles of *frequency*,

recency, and *vividness*. *Frequency* emphasizes the cumulative effect of repetition or repeated effort. Improvement in a given function is usually fairly rapid at first. Sooner or later, however, a "physiological limit" is reached, beyond which additional repetitions have little beneficial effect. This is as true of learning poetry or the multiplication table as it is of learning to skate or dance: repetition is necessary up to a certain point, but after that is probably of little value (p. 57).

Recency refers to the commonly observed fact that performances recently learned or practised are smoother and less subject to error than those unpractised for some time. Most busy adults have completely forgotten the algebra and Latin learned in high school or college (unless they happen to be teachers!); but they remember the necessary details of their business or profession. The negative side of the law of recency is covered by the law of disuse. Activities once learned but not recently practised tend to deteriorate with the passing of time.

Vividness or intensity means that active and interested exercise is more beneficial than passive and perfunctory repetition. If the learner sets out vigorously and with determination to learn a given task, if he has, as William James puts it, "the will to learn," he will progress more rapidly than if he simply sits passively, hoping that somehow the desired information will sink in. Vividness will explain why a boy learns his Scout Manual and the fine points of baseball more quickly than he learns to play the piano or recite his Sunday-school lesson.

All of these principles of learning are illustrated in modern advertising. Over and over again (frequency and recency) by newspaper, magazine, radio, and flashing electric sign (vividness), we are told about the merits of some particular cigarette or toothpaste. As a result of this bombardment most of us have become familiar with the names and slogans of many articles and "brands" and this very familiarity predisposes us to the purchase of these products. In the school room, these principles are constantly observed. There is drill and repetition (frequency); immediate review and recitation (recency); and finally, the attempt is made to relate the subject-matter of mathematics or literature in as many ways as possible to the everyday needs and interests of the student (vividness).

9. THE LAW OF EFFECT

The *law of effect* may be stated most simply, perhaps, as follows: When a response or series of responses leads to success or to a satisfying state of affairs, the connection between the situation and this response is strengthened, while other responses not so satisfying (i.e., annoying) are weakened and hence rendered less probable of recurrence. The law of

effect is really prior to the law of exercise, since it explains *how* the successful response came to be selected in the first place. This is clearly evident in the instances of animal learning which we have cited. When a cat, in learning to escape from a puzzle box, makes a variety of responses most of which are useless, one might expect from the law of exercise alone that *all* of these responses would be equally bound up with the situation "trying to escape from the puzzle box." Hence, all would tend to be repeated. Instead of this, however, the successful response, namely, pulling the loop or turning the button of the cage, soon gets the upper hand, while the useless responses drop out. Reference should be made again to Figures 11 and 13. Note how the escape of the cat or chick becomes faster and smoother as the animal goes more and more promptly to the proper exit. The successful response—that is to say, the one which leads to food or company—once having been made is soon enabled by exercise to supplant the unsuccessful reactions. Suppose that five cats have been taught to come to the call "kitty, kitty." Now suppose that later on, upon being called, cat no. 1 is given some fish and is petted; cat no. 2 is petted only; cat no. 3 is totally disregarded; cat no. 4 is sprinkled with water; and cat no. 5 is doused heavily with water. What will happen the next time these cats are called? Prediction is easy. Cat no. 1 will come quickly and promptly when called; cat no. 2 will probably come, but not so promptly; cat no. 3 may come, also, but still less promptly; cat no. 4 may try it once more, but if sprinkled again will surely give up; cat no. 5 will not only fail to come, but (if he is a normal cat) will run in the opposite direction. Here the end-effect—what happens to the cat—is the primary factor in modifying behavior.

10. PRESENT STATUS OF THE LAW OF EFFECT

Thorndike's law of effect has been criticized, modified, and (in some cases) rejected by experimental psychologists. One group of psychologists substitutes the concept of reinforcement for the law of effect. They argue that positive and negative reinforcement are less subjective and better defined "operationally" than are Thorndike's satisfiers and annoyers. If what happens *after* a response makes for continuation or repetition of the response, i.e., strengthens it, this positive reinforcement will appear in the learning curve. On the other hand, if what happens after a response makes for rejection or abandonment of this response, i.e., weakens it, this negative reinforcement will also appear in the learning curve. One need not assume that the effect is necessarily pleasant and satisfying or unpleasant and annoying. Reinforcements may, in fact, have *no* special emotional tone though they may be novel or striking or unusual, i.e., attention getting. Satisfaction or positive reinforcement has also been identified with need-

reduction (Hull)—a hungry dog eats and is no longer hungry, a timid child is commended by the teacher and no longer feels insecure.

Several systematic psychologists regard the law of effect as of secondary—if of any—importance. According to one view, the learner acquires new responses primarily by doing (Guthrie). Rewards or reinforcements are regarded as useful because they end the learner's search and make repetition of the last (successful) response more probable: but what is learned is what is associated, not what is rewarded. According to another view (Tolman), the law of effect is unnecessary because latent learning shows that behavior modification (learning) can take place without reward or other reinforcement. The concept of latent learning grew out of experiments in which rats were first placed in a maze and left to wander around without food reward. Later, when food was given at the end of a run, these rats learned the maze much more quickly than animals totally unfamiliar with the maze. Apparently the first group of rats had learned something from simply being in the maze, as there had been no positive reinforcement of their running behavior. The Gestalt psychologists who, as we have seen (p. 50), criticized Thorndike's animal experiments, regard the law of effect as having a rough descriptive value. They contend, however, that learning is far more organized, insightful and cognitive than is implied by Thorndike's laws.

Proper evaluation of the pros and cons of the large amount of experimental data bearing upon the law of effect would occupy far more space than can be given to it in this book. It must suffice, therefore, for us to summarize briefly the importance of the law of effect in learning theory. If satisfying states are taken to mean those which lead to positive approach-behavior, and annoying states those which predispose to avoidance and retreat, the law of effect is adequate to explain nearly all learning in a sensible way, i.e., to explain why some responses are retained and some are not. When the results of our behavior are pain, punishment, social disapproval, regret, and failure, the acts which lead to such annoying states are not repeated by the normal person. On the other hand, responses which bring food, comfort, kindly treatment, praise, and success will be repeated and through repetition (exercise) become more rapid and more certain. In those cases in which learning takes place although apparently not followed by satisfaction, but even by indifference or actual annoyance, it may be true that behavior of the opposite sort would be still more dissatisfying. The boy who dutifully practises his music lesson often has an ulterior motive which is not wholly inconsistent with the law of effect. And it is probable that the life of a martyr is not entirely devoid of psychological rewards.

It often happens, of course, that the end-result, i.e., the effect, while

providing self-satisfaction to the individual concerned, is actually injurious, or is socially disapproved or morally hurtful. Numerous examples could be brought forward to illustrate this situation. If a baby is fed or caressed whenever it cries; if the small boy is given money or candy whenever he whines or teases; and if the bully is applauded and feared by the other members of his group, such unfortunate behavior, because satisfying to the individual concerned, will tend to persist. To change or alter it, the end-effect must be made unsatisfying. Social disapproval, the withholding of something valued or desired by the child, mild punishment, or the substitution of new interests at once satisfying and more healthful, are common methods employed. Prizes, medals, the honor-roll, and other forms of approbation are some of the traditional ways in which the school has utilized the law of effect. The modern school has attempted to make the results of industry and application seem worth while in themselves to the child. To reach this end, it strives to get the desired responses and then through active and vigorous exercise to build them in. Modification of behavior is the fundamental problem in the psychology of learning, and such modification usually involves the principles of exercise and effect.

BIBLIOGRAPHY

1. Adams, D. K. Experimental studies of adaptive behavior in cats. *Comp. Psychol. Monog.*, 1929, 6, No. 1.
2. Allen, C. N. Individual differences in delayed reactions of infants. *Arch. Psychol.*, New York, 1931, 19, No. 127.
3. Birch, H. C. The relation of previous experience to insightful problem solving. *J. comp. Psychol.*, 1945, 38, 367-383.
4. Guthrie, E. R., & Horton, G. P. A study of the cat in the puzzle box. *Psychol. Bull.*, 1937, 34, 774.
5. Hilgard, E. R. *Theories of learning*. New York: Appleton-Century-Crofts, Inc., 1948.
6. Hilgard, E. R., & Marquis, D. G. *Conditioning and learning*. New York: Appleton-Century-Crofts, Inc., 1940.
7. Hunter, W. S. The delayed reaction in animals and children. *Behav. Monog.*, 1913, 2, No. 1.
8. Koffka, K. *The growth of the mind*. New York: Harcourt, Brace and Company, 1928, 180-193.
9. Köhler, W. *The mentality of apes*. New York: Harcourt, Brace and Company, 1925, 256-259.
10. Maier, N. R. F. Reasoning in white rats. *Comp. Psychol. Monog.*, 1929, 6, No. 3.
11. Maier, N. R. F. Reasoning in humans. I. On direction. *J. comp. Psychol.*, 1930, 10, 115-144.
12. Maier, N. R. F., & Schneirla, T. C. *Principles of animal psychology*. New York: McGraw-Hill Book Company, Inc., 1935, p. 449.
13. Ruger, H. A. An experimental study of the processes involved in the solution of mechanical puzzles. *Arch. Psychol.*, New York, 1910, 2, No. 15.

14. Shellow, S. M. Individual differences in incidental memory. *Arch. Psychol.*, New York, 1923, 10, No. 64.
15. Thorndike, E. L. *Animal intelligence*. New York: The Macmillan Company, 1911.
16. Thorndike, E. L. *Human learning*. New York: Appleton-Century-Crofts, Inc., 1931, 3-29.
17. Warden, C. J. *A short outline of comparative psychology*. New York: W. W. Norton and Company, 1927.
18. Warden, C. J., & Warner, L. H. The sensory capacities and intelligence of dogs. *Quart. Rev. Biol.*, 1928, 3, 1-28.
19. Warden, C. J., & Jackson, T. A. Imitative behavior in the rhesus monkey. *J. genet. Psychol.*, 1935, 46, 103-125.
20. Washburn, M. F. *The animal mind*. New York: The Macmillan Company, 1936, 370-372.
21. Woodworth, R. S. *Contemporary schools of psychology*, rev. ed. New York: The Ronald Press Company, 1948.
22. Yerkes, R. M. Modes of behavioral adaptation in chimpanzee to multiple-choice problems. *Comp. Psychol. Monog.*, 1934, 10, No. 1.

4

KÖHLER'S EXPERIMENTS IN PERCEPTION AND LEARNING AND THEIR IMPORTANCE FOR GESTALT PSYCHOLOGY

1. THE MEANING OF GESTALT PSYCHOLOGY

THE POINT of view and method of interpretation represented under the name Gestalt psychology, had its rise in Germany as recently as 1912. Max Wertheimer is generally credited with being the founder of the movement in Germany, but Kurt Koffka, formerly of the University of Giessen, and Wolfgang Köhler and Kurt Lewin, formerly of Berlin, through academic appointments in this country have become the principal champions of the Gestalt viewpoint in America.

Perhaps the simplest way of describing just what this school of psychology stands for is to indicate first what it is not. "The psychologist," says Titchener, "seeks, first of all, to analyze mental experience ('consciousness') into its simplest components." "The rule, or measuring rod, which the behaviorist puts in front of him," says Watson, "always is: Can I describe this bit of behavior I see in terms of stimulus and response?" To both of these programs, Gestalt psychologists stand strongly opposed. It is bad psychology, they say, to contend that the complex behavior of men or of animals can be explained genetically as an accumulation of specific S-R bonds—the so-called "bundle hypothesis." And it is equally invalid, they hold, to apply strict analysis to complex sensory data with the expectation of finding some fundamental psychological "atom" out of which experience is built. To the Gestalt psychologist sensory elements appear only after careful and somewhat "unnatural" introspection. The *real* data of experience are organized and extended wholes, never mosaics; we do not encounter specific elements either in consciousness or in behavior. And so instead of the world being to the infant "a big, blooming, buzzing confusion" (William James) from which here and there bits must be laboriously picked out and tied together by the process of association, there is—according to this view—a certain degree of orderly arrangement in sensory data to which even the very young child may respond without previous

learning. Later on adults react not to specific stimuli but to the pattern or total organization of objects around them. These Gestalten or configurations, as the word has been translated, are the true "mental elements." For this reason, Gestalt psychology studies the organization of these unitary experiences; how they come about; what "laws" govern their changes; and upon what factors they depend.

2. STUDIES OF VISUAL PERCEPTION

The first problem attacked by the Gestalt psychologists was that of how we perceive apparent movement. All of us know that when pictures which differ slightly follow each other in rapid succession we see *movement*—not a series of separate and distinct "stills." This common observation was the basis of Wertheimer's classical experiments on apparent movement. The principle of these experiments will be made clearer by an illustration (Fig. 16). Suppose that through a vertical aperture light is thrown at short

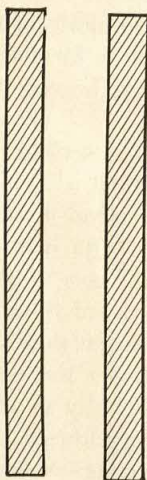
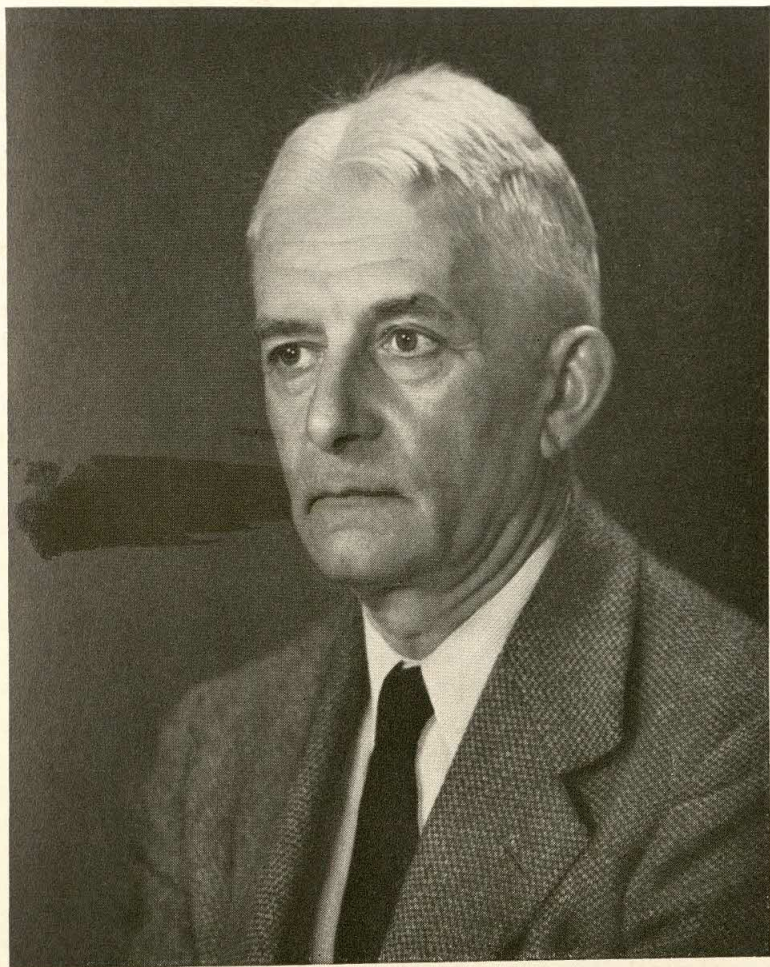


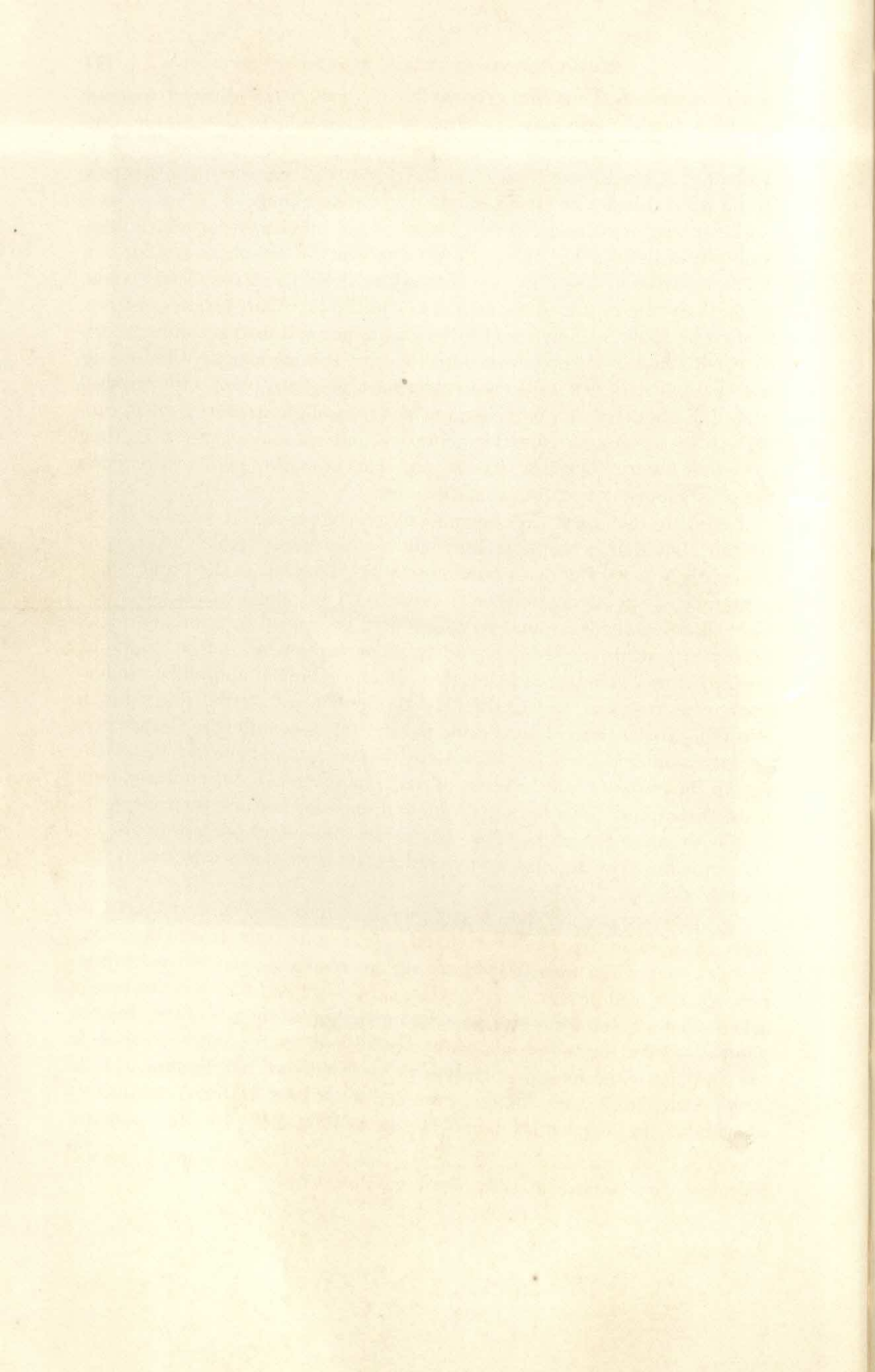
FIG. 16
TO ILLUSTRATE THE PHI-PHENOMENON

regular intervals upon a screen in a darkened room, thus producing a series of flashing white lines against a black background. Now, if a second beam of light is projected through a second aperture, slightly to the *right* of the first, and somewhat later in time, logically it would seem, we should see first one and then the other of *two* parallel lines, the one to the right of the other. When the time interval between the first line and the second line is relatively *long*, this is exactly what we do see. But a very different result



WOLFGANG KÖHLER

1887-



may be obtained. If the timing is carefully adjusted, the observer sees not *two* lines, but a single line oscillating or "jumping" from side to side. The addition of the second line changes greatly the stimulus-value of the first. Instead of a simple addition of stimuli, what we perceive is a complex fusion which cannot be analyzed into two separate lines.

The appearance (under these conditions) of movement for which there is no physical basis was called by Wertheimer the *phi*-phenomenon.¹ *Phi* illustrates what is meant by perceptual Gestalten. From the Gestalt viewpoint, the value of this phenomenon lies in the fact that it shows our perception of apparent movement to be an unique and unanalyzable experience, not a rubber-stamp impression of nature. It is evident, say the Gestaltists, that our experiences do not correspond point-for-point with physical stimuli. Rather they fall into coherent and meaningful patterns which cannot readily be separated into elementary sensations and images. According to Gestalt theory, therefore, sensory elements are not the units with which the psychologists can most profitably work.

It is likely that most experimental psychologists would agree with the Gestalt view that perceptions are "whole experiences" which are not the sum of their parts. But many would be inclined to explain the stability and constancy of our perceptions of an object (its meaning) by the variety of associations built up around it, rather than by appeal to principles of inherent organization. Thus, one's visual perception of a rose might be attributed to a complex of associated olfactory, tactual and other stimulations, plus personal and emotional experiences and verbal descriptions. With this explanation of perception in terms of association the Gestalt psychologists would not concur. They deny that objects and events have meaning for the observer only in terms of past experience. On the contrary, they insist that stimuli have form, pattern, and meaning because their elements "come" organized. And they point to the fact that except perhaps in disease and some forms of insanity, no psychological experience is ever totally unorganized.

This inherent organization of experience is illustrated, Köhler thinks, in those cases wherein persons born blind later acquire sight. It is to be noted, he says, that while such individuals do not recognize geometrical forms, e.g., squares and circles, *as such*, they do *understand* the question when asked what a given form is. This must mean, according to Köhler, that organizations and arrangements of the visual field are present even though the stimulus situations are not yet tagged with verbal symbols (with names). An adult born blind, whose sight has been restored, is able to identify a coin lying on his palm. He may call the coin a quarter, perhaps

¹ Wertheimer's experiments on the *phi*-phenomenon have been repeated by several investigators and his essential results have been substantiated.

inaccurately, but he distinguishes the coin as being something apart from his hand.

The visual field is chaotic, relatively speaking, when there are serious brain lesions or diseased conditions in the optical centers of the brain, so that only separate bits can be grasped visually at one time. Köhler cites a study made by Gelb and Goldstein of a patient suffering from brain injury. This man had learned to rely upon motor and kinesthetic experiences to piece out his visual data. When his name was written in large letters, and presented to him, he would perceive the first few letters as a group and then guess the remainder from the context and from the amount of movement needed to span the whole name. But if several lines were drawn across his name, as shown in Figure 17, the patient could no longer see the letters as a *name*. Since he did not see his name as an organized perceptual unit, the lines drawn in over the letters were mixed indiscriminately with the letters. What the patient saw was a single confused pattern—not two separate and independent groupings, the one of lines and the other of letters. Nothing was added from other parts of the field to give the letters complete organization and in consequence no Gestalt (name) appeared.

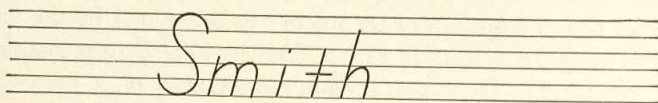


FIG. 17

TO ILLUSTRATE THE EXPERIMENT OF GELB AND GOLDSTEIN

Perhaps it should be said at this point that while some primitive form-sense (inborn perceptual Gestalts) *may* exist apart from experience, such an assumption is not necessary to explain the facts of perception. The retina furnishes an area upon which objects are spread out in two dimensions. To the spatial cues from this source, other sensory experiences add associations which lend meaning to our perceptions. Experiments have shown that learning is at least as important as native factors in judgment of distance and depth (see Chapter 14). And learning certainly gives meaning and value to our understanding of the objects themselves. Basic "native" Gestalts may, of course, exist. But they must still be accounted as hypothetical rather than as proven.

3. SOME PRINCIPLES GOVERNING THE FORMATION OF PERCEPTUAL GESTALTS

a) *Figure-Ground Relations*

To the Gestalt psychologist, the distinction between "figure" and

"ground" is decidedly important to the psychology of perception. Every Gestalt exists as a figure *against* a more general, and usually vaguer, background. In auditory perception, for example, a melody is heard against a background of silence or against a vague mixture of miscellaneous noises. And in the same way, every visual configuration, a sunset or a prizefight, is projected against a larger and less clearly defined background.

The figure-ground relationship has been carefully studied in the sphere of visual perception. Ambiguous designs in which figure and ground are interchangeable serve to define most clearly the distinction between these two factors. Figure 18 shows an ambiguous geometrical construction. Note

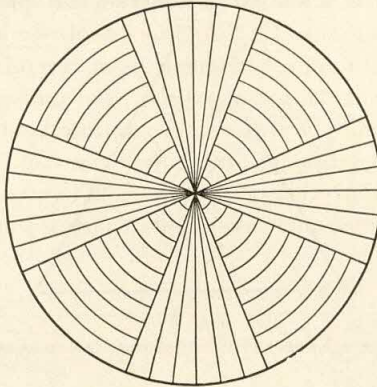


FIG. 18

AN AMBIGUOUS DESIGN TO ILLUSTRATE FIGURE-GROUND RELATIONSHIP
(From Koffka.)

that the four-pointed "cross," the one composed of radii, "stands out" rather easily as a Gestalt against the concentric circles which form the ground. Now interchange figure and ground by fixating intently upon the circles, and note that the arcs in the *new* figure, again a four-pointed "cross," are *true* arcs, i.e., cut-off sections of circumferences, whereas before they were visible parts of unbroken circles. The circles behave differently depending upon whether they are in the figure or in the ground. Following the lead of Rubin, Gestalt psychologists have developed other characteristic differences which they believe serve to distinguish figure and ground. The figure is more clearly formed than the ground (i.e., has better structure) and is more resistant to change; is more vivid; holds its color better; is more solid and more substantial.

Koffka illustrates the principles of figure-ground relationship in the following way. Look, he says, through a stereoscope at a slide which to the one eye presents a solid blue, and to the other a circle made up of alternate blue and yellow sectors. The yellow "slices" easily stand out against the

blue ground in the single figure which results; but the blue sectors have a hard time of it as "figure" against a yellow background, being overwhelmed by the surrounding blue.

b) Closure: the Law of "Pregnance"

The drawing in Figure 19 illustrates another principle stressed by the Gestalt psychologists, namely, that of "closure." The figure is seen at once to be a triangle, although it has only two true angles. Our drawing, to be sure, is not closed; but the direction that closure will take is so clearly indicated that reaction is immediate and undisturbed. Had this figure been exposed for a fraction of a second on a screen, the opening would not have

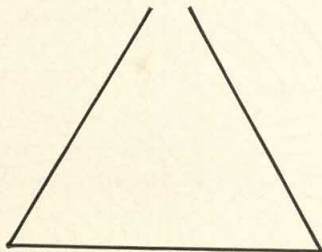


FIG. 19

THE FIGURE IS SEEN AS A COMPLETED TRIANGLE, AN ILLUSTRATION OF CLOSURE.

been perceived at all. Another illustration of the phenomenon of closure is given in Figure 20. If one gazes intently at the staircase it suddenly turns upside down and becomes a staircase viewed from underneath or a piece of overhanging masonry. When the "new" figure leaps into view closure is said to have taken place. The puzzle pictures for children in which one searches through a maze of lines for a man's face (which suddenly appears) illustrate the same phenomenon. The meaningful figure, when it emerges, is seen as a totality and not piecemeal. In the interval during which one is seeking for a solution—before closure—there is decided discomfort and even tension and nervousness.

The explanation of closure given by the Gestalt psychologists is that, first of all, there is an inherent tendency toward "form" and meaning in all of our perceptions. Furthermore, this tendency is so strong that when the external situation is incomplete or confused we strive to complete it and are dissatisfied until we do. Closure is a special case of the law of *Prägnanz* (usually translated "pregnance") or principle of equilibrium. According to this law, every experience (perceptual or otherwise) tends to complete itself and to become as "good as possible." Thus, our perceptions of the drawings in Figures 19 and 20 exhibit a decided tendency to arrive

at a condition of maximum simplicity. Closure and the law of Prägnanz are used by Gestalt psychologists as explanatory principles in learning as well

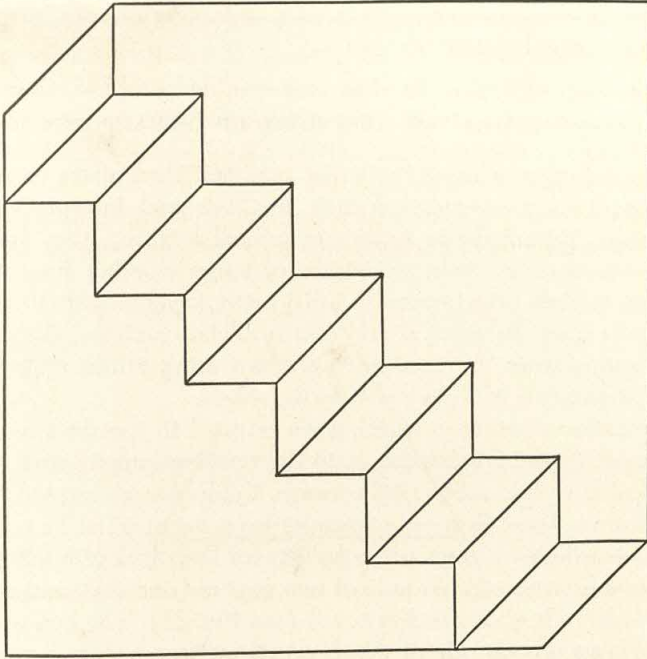


FIG. 20

THE AMBIGUOUS STAIRCASE FIGURE, ILLUSTRATING CLOSURE

as in perception. Thus, the solution of a problem or the attainment of a goal brings closure and ends the active search of the learner.

c) Other Principles of Organization

Several conditions are important to Gestalt psychology in the organization of the visual field. These may be outlined briefly:

(1) *Similarity*. In a total stimulus-situation, things which are alike tend to fall "naturally" into groups or to be perceived as wholes. Dots of the same size and shape fall into patterns; a face recalls a friend's; men in uniform fall into groups or squads.

(2) *Proximity*. Things and objects close together in space (or time) tend to be grouped together as units. Illustrations are sounds which follow each other closely; parallel lines, which when close together are seen as fences.

(3) *Closed forms and "good" contours*. Closed geometric forms (circles or squares, say) are seen as units. Also, lines or forms which are stable

(balanced or symmetrical architectural forms) are more readily perceived as unitary than are vague and poorly made figures.

These principles are believed to hold generally no matter what the experiences of the perceiver. They represent facts which are "given" or are inherent in the perception.

4. LEARNING FROM THE VIEWPOINT OF GESTALT PSYCHOLOGY

The Gestalt psychologists have not confined themselves to studies of perception, but have extended their methods and interpretations into other areas. Particularly in studies of animal learning, they have many experiments to their credit which are of value whether interpreted according to Gestalt principles or not. To many psychologists these experiments seem more substantial and less artificial, perhaps, than much of the perception work. As illustrative of the Gestalt attack upon learning, several experiments of Köhler will be described.

The question whether in learning we respond to specific stimuli in the environment (strict S-R bonds) or to the *relations* among such stimuli is of practical as well as theoretical concern. Köhler has attempted to answer this question in the following interesting experiment. First he trained two hens (birds which are generally notorious for their lack of intelligence) to expect food from the darker only of two papers (one dark and one light) glued side by side on a wooden board (see Fig. 21). The hen was placed in a wire coop on one side of which apertures were so arranged that the fowl could easily thrust its head out to peck grain from the paper-covered board next to the coop. Grain was scattered in equal amounts upon both papers. Whenever the hen pecked at the grain upon the dark or positive paper she was allowed to eat, but whenever she pecked at the grain upon the light or negative paper, she was at once shooed away. After 400 to 600 trials in which the positions of the positive and negative papers were frequently interchanged to prevent any association by the fowl of the positive stimulus with the right or left side, the hen at last learned to select the positive and to avoid the negative stimulus with great regularity. This part of the experiment constitutes what is called the training series. In the critical tests which followed Köhler kept the positive paper on the board, but substituted for the negative another paper still darker than the positive stimulus which the hen had been trained to select. The question now was whether the hen would continue to peck at the *specific* paper from which she had been laboriously trained to eat, or whether she would respond to the brightness *relation* and select the still darker but neutral paper to which she had never gone. Köhler's results seem fairly conclusive. Four hens, two trained to go to the brighter of the two papers and two trained to

go to the darker of the two, selected the neutral paper or new stimulus in about 70 per cent of the trials and the original positive stimulus in 30 per cent. These results show, Köhler thinks, that there are in nature simple organizations of brightness differences which are extremely primitive, to be sure, but yet fundamental. Even so stupid a creature as a hen will respond to the "darker than" relation, rather than to the specific stimulus which she has been carefully trained to select.

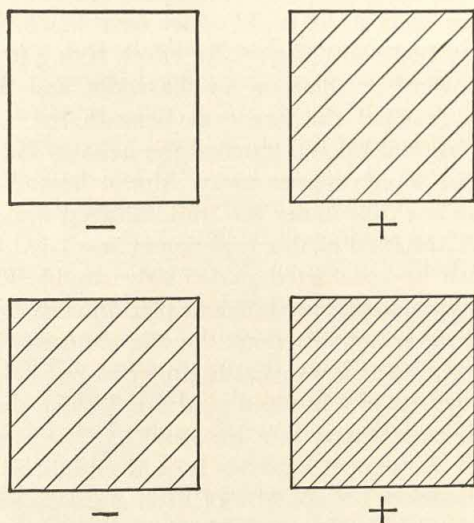


FIG. 21
TO ILLUSTRATE KÖHLER'S EXPERIMENT

Köhler repeated this experiment with chimpanzees and also with a three-year-old child. The child was allowed to learn by the method of trial-and-error that only the brighter of two colored boxes contained candy. Forty-five trials were necessary in order to establish this differential response to the point where no errors were made, i.e., the brighter was always chosen. In the critical tests which followed, a new and still brighter box was presented with the old positive stimulus, the negative darker box now being removed. Köhler reports that in this series the child "invariably" took the new bright box, instead of the one in which it had formerly found candy. This again is good evidence, he says, that human beings react toward situations as *related wholes*, and not to bits or portions of the environment.

Köhler's ape experiments—which are by all odds his best known—were conducted during the years 1913–1917, upon the small island of Tenerife, located in the Canary Islands, a short distance off the African coast. Here

a colony of chimpanzees was established, and owing to the isolated and tropical nature of the country, experiments could be carried on under conditions closely resembling those of the animal's native habitat. In one experiment a chimpanzee was not fed in the morning as usual, but instead his food was fastened to the roof of his cage and a box thrown casually upon the floor of the cage some distance from the point where the fruit hung suspended. The ape had never used a box as an implement, and hence ignored it completely. By pulling the box under the fruit, however, and climbing upon it, he could easily have reached the fruit. The animal spent many hours in unsuccessful effort, trying to reach the fruit by jumping up toward it, climbing up the walls, and the like. Finally the experimenter dragged the box over beneath the suspended fruit, stepped upon it, reached up and touched the banana. He then got down and placed the box some distance away. Almost immediately the chimpanzee pulled the box over under the fruit, climbed upon it, and pulled down the food. A variation of this experiment was tried by Köhler with another ape which he considered particularly stupid. This animal had often seen other chimpanzees use boxes as platforms from which to reach food, but he had never actually done the trick himself. To see whether he had learned from his fellows what to do when put into a situation requiring the employment of a box as a platform, Köhler fastened a banana to the roof of his cage, threw in a box, and left the animal to its own devices. The ape's subsequent behavior is quite enlightening as a demonstration of how hopeless the solution of even a simple problem may be unless the relations involved are clearly apprehended. This monkey at once ran to the box, but instead of dragging it under the fruit, he either climbed upon it and jumped up, or else climbed upon the box, hopped down, and quickly ran over to jump up from the ground under the fruit.

This experiment, says Köhler, shows how necessary for learning it is to see the situation involving the task as a whole or in its entirety. The first ape did not connect box and fruit until the relation was demonstrated to him by the experimenter. Once the relationship was seen, the box ceased to be simply a box and became an implement—something to be used in getting food. More technically speaking, at the instant this relation was seen, closure occurred. The stupid ape, it appears, knew that the box and jumping up were both involved in the task of getting the fruit, but that was as far as his analysis went. Had the connections ground-box-fruit been made in their proper sequence, the situation would have been organized at once, and the grouping would have become meaningful. But this last Gestalt was never formed by the stupid ape; instead there were for him two separate groupings, box-jumping-up and jumping-up-under-fruit.

5. OTHER EXPERIMENTS ON LEARNING AND MOTIVATION

In another series of experiments Köhler attempted to see whether Sultan, apparently his most intelligent chimpanzee, could combine two sticks into one useful implement. The sticks were two hollow bamboo rods, the one being enough smaller than the other so that it could be fitted easily into the end of the larger to form a single long stick (see Fig. 22). All of



FIG. 22

A CHIMPANZEE FITTING A SMALL STICK INTO THE BORE OF A LARGER ONE
(Adapted from Köhler's *The Mentality of Apes*.)

the animals had frequently used single sticks to pull in bananas and other fruit placed outside of the bars of their cages. But none had been given the task of joining two sticks into a single one and using the resulting long stick as an implement with which to pull in fruit.

The setup of the experiment was as follows. The chimpanzee was put into a cage in which were two sticks, and several pieces of fruit were placed outside of the bars too far away to be reached by either stick alone, but within easy reach of the joined stick. Judged in the light of this rather simple situation, the animal's behavior seems incredibly stupid to

human eyes. First, he tried to reach the fruit with one stick and then with the other. This failing, he next pushed one stick out as far as possible, and then with the second stick pushed the first one on until it finally touched the fruit. This actual contact, Köhler notes, seemed to give the animal great satisfaction, but it didn't give him the fruit! At last, as the animal seemed no closer to a solution than at the start, the experimenter gave him a hint by sticking one finger into the opening of the larger stick directly before his eyes. But this cue failed to help any, and after an hour or so of futile effort, the ape apparently lost interest and gave up the task as hopeless. He continued to play with the two sticks, however, and after some manipulation, holding the one stick in the left hand and the other in the right, he accidentally got them together as shown in Figure 22. The first connections were loose, so that the sticks frequently fell apart, but the animal persisted, with great eagerness pulling in not only all of the fruit, but all other small moveable objects, such as stones and sticks, within reaching distance. On the following day, after some desultory pushing of the one stick by the other (repetition of the old useless behavior), Sultan quickly joined the two and got the fruit.

This experiment, as interpreted by Gestalt principles, shows again the growth of elements, at first unconnected, into an organized whole. Until the ape perceived the two sticks as capable of forming a single unit in the situation, reaching-the-food, no Gestalt appeared. When the two sticks became one implement, however, a new pattern was at once formed and closure took place. Many illustrations of the growth of such patterns are cited by Köhler. A box will be employed as a platform for reaching fruit if it is in close proximity (p. 72) to the fruit; but not if it is in a far corner of the cage. A stick will be used to pull in fruit from the outside of the bars if it is close enough to where the fruit is lying to "get into the same picture" with it, but not if it is at some distance, or is only potentially present as in the branch of a tree which has been thrown into the cage.

The crux of the matter, according to the Gestaltists, is that, to be used as an implement, the box or the stick must exist in an organized visual field, the relations of whose parts are perceived. Learning, according to this view, is not then a bundle of specifically acquired stimulus-response bonds. When the relations involved are quickly grasped, i.e., when the animal almost at once gets the whole picture of connections and mutual implications, as, for instance, climb-on-box-to-get-fruit, it is said to have "insight." A situation is experienced without insight when its various elements are seen as parts having no dependence or orderly arrangement. Insight into a problem, then, involves seeing its various implications as all of one piece.

A series of investigations in which Gestalt principles are applied to the

general field of motivation has been carried out by Kurt Lewin and his students. Lewin, whose work has been principally in child psychology, has made many contributions both to the theoretical and experimental aspects of Gestalt theory. In one study made under Lewin's direction, the difference between the retention of finished and unfinished tasks was studied. Nearly 200 subjects, including both adults and children, were given from eighteen to twenty-two fairly simple problems such as solving match-stick puzzles, writing names of cities beginning with *L* and the like. In one half of the tasks, the subjects were interrupted by the experimenter before completion of the task; in the other half they were allowed to finish. Later, the subjects were required to write down all of the tasks which they could recall. If the interrupted and completed tasks had been retained about equally well, the ratio IR/CR (where IR = number of interrupted tasks, and CR = number of completed tasks) would closely approximate 1.00. Actually, the ratios for the different groups of subjects were nearer 2.00 than 1.00—showing the unfinished tasks to be much more clearly remembered than the finished. In other variations of the experiment, the subjects were interrupted at different stages in their solutions. Retention of tasks, broken toward the end or in the middle, was much greater than retention of tasks interrupted at or near the beginning.

The greater retention of unfinished than of finished tasks is explained as an instance of the law of pregnancy. The finished task is a completed Gestalt—the goal has been attained, tensions relieved, and relaxation obtained. The unfinished task, on the other hand, leaves the individual unsatisfied and often irritated. No closure has taken place. The “need” to complete a task once begun is shown, according to Gestalt theory, by the subject's disturbance on being interrupted and his willingness to resume as soon as opportunity is provided.

To the non-Gestalt psychologist the greater vividness of the unfinished tasks in recall may be an illustration of the tendency shown by many activities—unfinished as well as finished—to “perseverate” after removal of the stimulus. Catchy tunes tend to run in the head; interesting happenings, unsolved or disturbing problems, spontaneously recur over and over again. Physiological activity often continues for a considerable period of time *after* the subject's active response has ceased. Such activity may be greater for unfinished than for finished tasks.

Besides building up a thoroughgoing experimental basis for their theoretical concepts, Gestalt psychologists have advanced cogent and searching criticism of other methods and points of view. Köhler's criticism of the methods employed in many animal learning experiments is noteworthy. In learning problems of the “choice reaction” type, the animal

is trained to go to one stimulus, a red light, say, and to avoid another stimulus, a green light, by being fed when it chooses the one and shocked when it takes the other. In such experiments, Köhler points out, instead of giving the animal an electric shock in the feet when it responds to the negative stimulus it would be much more logical from the point of view of the animal's own experience to have the negative stimulus do something to the animal—move forward or frighten it in some way. Not only would this be closer to the natural situation for the animal, but the connection between reward and the one stimulus, and punishment and the other, would be much more obvious and sensible. The latter situation, Köhler thinks, would furnish a good Gestalt; the former, which is traditional in animal experimentation, a poor one, since it is well-nigh impossible for the animal to organize into one situation punishment-in-feet-for-taking-the-wrong-light. As noted in Chapter 3, pp. 50–51, both Köhler and Koffka have criticized the animal experiments of Thorndike on much these same grounds. The animals in these experiments, they say, were forced to learn by trial-and-error, since there was no possible way—in the light of their experience—by which they could organize such acts as opening a latch or pulling a loop into the situation getting-out-of-a-box-and-getting-food. Much of the so-called stupid learning of animals, Köhler believes, results when almost impossible tasks are set for the animal, tasks which cannot readily be organized into a meaningful pattern, if they can be so organized at all.

6. STUDIES IN GESTALT THEORY IN AMERICA

The work of the Gestalt psychologists, most of which had been done in Germany, inspired American psychologists to undertake investigations to verify or to extend the German work. Much of what has been done in the field of perception and learning is technical and difficult to describe simply or in brief space. Several studies will serve however as illustrations of the kind of work being done. De Silva has investigated the factors determining apparent visual movement such as is represented by Wertheimer's *phi*-phenomenon (p. 64). Gestaltists argue that the apparent movement which occurs in the *phi* is a clear-cut demonstration that responses are non-additive. Actual movements and illusions of movements, they hold, are fundamentally the same, the latter being limiting cases of the former. To some extent De Silva substantiates this viewpoint. The perception of real or actual movement, however, he found to depend upon speed or velocity; while the perception of apparent movement was conditioned by the separation in space of the stimuli and by their degree of intensity. De Silva especially emphasizes the subjective factors present in the percep-

tion of apparent movement. Hulin has studied apparent tactual movement by applying two pressure points upon the skin of the forearm. These tactual stimuli were applied sometimes together and sometimes as much as .3 sec. apart; spatially, they were separated in eight steps from 5 to 150 mms. Hulin's subjects were instructed to report when they experienced apparent movement from one point to the other. About 30 per cent of the 13,500 judgments given by seven observers indicated the experience of apparent movement, i.e., configurations in the tactual field. Apparently, tactual *phi* are dependent upon definite temporal and spatial relations among the stimuli. Not all tactual Gestalts, of course, need be illustrations of the *phi*-phenomenon. *Phi* is somewhat harder to observe in the cutaneous than in the visual sense, but it is found there.

An interesting investigation, growing out of Köhler's work with chimpanzees, is the study by A. Alpert of the rôle of insight in the learning of very young children. Forty-four nursery school children ranging in age from nineteen to forty-nine months were subjects. Situations closely resembling the problems set by Köhler for his apes were devised. For example, a toy was placed on a shelf out of reach and a box or chair left near by; if the child used the box or chair as an implement—stood upon it—the toy could be secured. Again a toy would be placed outside of the bars of a playpen and a stick left handy near by with which the desired object could be fished in. Sometimes this problem was complicated by providing two sticks and a toy—a short stick, which was too short itself to reach the toy, but which could be used in raking in a long stick, which in turn could then be used to fish in the toy. The question in these experiments was whether the child would show the same sudden insight into the task as was exhibited by several of Köhler's apes. Alpert discovered what she called "immediate insight"—a quick, sure solution—more often than Köhler found it for his apes. In many cases, however, insight came gradually or partially, preceded by what seemed to be a certain amount of fumbling trial-and-error. It is impossible to say just how much tentative trial-and-error preceded this "immediate insight," but on the whole it was probably more often present than not. If comparisons between species are valid—which is doubtful—the young children were more insightful in their solutions than the chimpanzees.

Köhler's experiment on relational learning in hens, described on p. 71, has been repeated by Taylor with conflicting results. Köhler was interested, as we have indicated, in refuting Thorndike's theory that learning consists in the establishment of specific S-R bonds; and in showing instead that the hens responded to the relationship "darker than." Taylor trained his hens using black, grey, and white surfaces; and to speed learning, gave an electric shock to the hens that pecked from the "wrong"

surface. In his critical test, 82 per cent of the responses were to the *specific* shade to which the hen had been trained to respond positively—not to the brightness relation. Taylor suggests that Köhler's hens may not have been sufficiently trained to guarantee responses to the specific stimulus; and that it is a question whether hens can really distinguish the slight brightness differences which Köhler used. In the light of Taylor's experiment, Köhler's results appear to be less conclusive than he thought. It is only fair to say, however, that the transpositional type of learning exhibited by Köhler's hens has been demonstrated with other animals and for a variety of stimuli.

7. PRESENT STATUS OF GESTALT PSYCHOLOGY

Gestalt psychology has been presented by its champions with vigor and enthusiasm, but it has not won many outright adherents in America. Objection has been raised to its theoretical assumptions rather than to its experimental procedures. It has been urged, for instance, that the term Gestalt conveys nothing essentially new. Perception has long been regarded by non-Gestalt psychologists as a unified response or as a "combining activity." Few would care to argue that perception is a mechanical photographic registration; or an accumulation of specific S-R bonds "bundled" together into a meaningful whole. When the Gestaltists attack the mechanical theory of learning or the "bundle hypothesis" they are in reality demolishing a not very strongly defended point of view.

Non-Gestaltists agree that Gestalt psychologists have contributed notably to our better understanding of perception. But they hold that their contribution has been in the nature of a refinement of understanding and a better distribution of emphasis than in the discovery of essentially new facts. The distinction between figure and ground, for instance, is not especially original. The well-known phenomena exhibited in binocular rivalry² and in the ambiguous figures show how a detail can stand out against a background at one instant and become a part of it later on. These have been recognized for a long time. It should be said, however, that the Gestaltists have carried the study of the figure-ground relation and the laws governing it much further than have other workers.

In the field of learning, criticism has been leveled at the term *insight*. The contention is that this word while descriptive adds nothing in the way of fundamental explanation. Apparently there are two ways of de-

² Binocular rivalry occurs when two dissimilar pictures are placed in a stereoscope. The result is a competition or rivalry between the two, first the one and then the other picture being suppressed.

fining insight. One is to make the term purely descriptive of some end-result. A child operates a new mechanical toy quickly and deftly, with clear understanding of its function, and we say he shows insight. Used in this way, the term is neat and unambiguous and would be objected to by no one.³ The second definition makes insight an explanatory principle like trial-and-error or the conditioned response—a person learns “by insight.” The Gestaltists seem often to mix these usages, sometimes employing the term descriptively, sometimes as an explanatory principle. It is the latter usage which non-Gestaltists find it hard to accept, since there is a certain mystical or intuitive flavor about it.

Insight, according to Gestalt psychology, appears when a man or an animal fully grasps the principles involved in a task or sees the proper relationships forthwith. It has been picturesquely called a “click” or “A-ha” experience. But as indicated elsewhere, it is very difficult to know just how much trial-and-error learning has preceded the flash of understanding. It may be that this is always the case, provided the learning is closely analyzed. In hard and little-comprehended tasks, we know that both men and animals resort to hit-or-miss learning until “by chance” the successful move or act is performed. In simpler problems there is often a marked reduction of random activity, and sometimes the sudden appearance of insight. This happens also when the individual is able to bring to bear information or knowledge from previous and somewhat similar situations (p. 55). It would seem plausible, then, to regard insight, genetically at least, as the end-result of a much reduced trial-and-error period. The previous tentative trials exist, to be sure, but are not represented in the particular learning curve under scrutiny, prior to the appearance of insight. It has been pointed out above (p. 51) that the insight shown by Köhler’s apes was in part at least a function of the tasks set the animals. When the learner is well acquainted with the facts of the situation so that a solution (reorganization) comes swiftly, insight would seem to be the correct term to apply.

Criticism has sometimes been directed against Köhler’s animal work on the score that the experimenter’s presence on various occasions during an experiment constituted an uncontrolled factor. Animals are quick to pick up cues from an entirely innocent experimenter, as has often been shown in the cases of horses and dogs which calculate or perform other supposedly “intellectual” feats that turn out upon investigation to be anything but “intellectual” (see p. 54). Again, the results of many Gestalt experiments are not statistically well established. In the experiment quoted on page 71, for instance, the hens responded to the relation of

³ Note that in psychiatry a person is said to possess insight when he realizes that his behavior is abnormal. An insane man lacks insight—at least part of the time.

brightness in 70 per cent of the trials. Although this result makes the relational response more probable than the specifically learned one, it is not sufficient to establish it with certainty.

8. INSIGHT AND EDUCATIONAL PRACTICE

On the whole, the doctrine of insight, with its emphasis upon understanding and generalization, has had a salutary effect upon pedagogy and educational theory. Thorndike's final formulation of the law of exercise, for example (p. 57), minimized the value of exercise per se as an important factor in learning. Less stress is now being placed upon a strict frequency theory of learning and upon the all-importance of simple drill and mechanical repetition. Emphasis upon insight has been helpful, also, in leading teachers to strive for the emergence of broad principles with a consequent better transfer to other situations.

The notion of insight fares better as a broad generalization than as a guiding principle in practical learning situations. Suppose that a class in arithmetic is being taught long division. What advice, exactly, would the Gestalt psychologist give the teacher? According to Wheeler, a leading American exponent of Gestalt psychology, an act is learned "when the stimulus-pattern in which the goal figured presented relationships on the learner's level of insight, that is, fitted his level of maturation (ref. 21, p. 317). Presumably, the teacher would be told to set a goal—praise, an honor-roll, a promise of more interesting tasks—and then wait patiently for the child's insight to develop. Insight might be expected if the child possessed the requisite physical and mental maturity. Unless he does, of course, learning would be slow and halting in any case. And even if he were mature enough, his insight would surely become more probable if the teacher employed the time-honored principles of repetition, explanation, and illustration. In other words insight should be given an opportunity to appear; it cannot be expected to spring forth without effort on the part of the learner.

This has been well shown in several experiments by Katona, in which memorizing by rote was compared with memorizing in which a better understanding of the task was afforded. This last was achieved by providing varied experiences which would acquaint the learner with the principles involved in the problems. Katona found that learning with understanding improved retention and enabled the learner to perform new tasks better. Insight in these experiments is quite similar to transfer of training through "generalized experience" (p. 95).

In the preceding sections we have sketched in brief outline the point of view and some of the experimental findings of the Gestalt psychologists and have given, as well, some of the more obvious criticisms which have been directed against their claims. Whether psychologists concur with the Gestalter's viewpoint or not, most of them are agreed as to the intrinsic value of their experimental work. As a fresh and vigorous movement with emphasis upon experiment, Gestalt psychology has played an important rôle in the development of perception and learning.

BIBLIOGRAPHY

1. Alpert, A. The solving of problem situations by pre-school children: an analysis. *Teachers Coll. Contrib. to Educ.*, 1928, No. 323.
2. Dashiell, J. F. A survey and synthesis of learning theories. *Psychol. Bull.*, 1935, 32, 261-305.
3. DeSilva, H. R. An experimental investigation of the determinants of apparent visual movement. *Amer. J. Psychol.*, 1926, 37, 469-501.
4. Ellis, W. D. A source book of Gestalt psychology. New York: Harcourt, Brace and Company, 1938.
5. Guilford, J. P., & Helson, H. Eye-movements and the phi-phenomenon. *Amer. J. Psychol.*, 1929, 41, 595-606.
6. Hartmann, G. W. Gestalt psychology. New York: The Ronald Press Company, 1935.
7. Hulin, W. S. An experimental study of apparent tactual movement. *J. exper. Psychol.*, 1927, 10, 293-320.
8. Hunter, W. S. Experimental studies of learning. In Murchison, C. (Ed.), *Handbook of general experimental psychology*. Worcester, Massachusetts: Clark University Press, 1934, 509-514.
9. Katona, G. Organizing and memorizing: studies in the psychology of learning and teaching. New York: Columbia University Press, 1940.
10. Katz, D. Gestalt psychology: its nature and significance (translated by Tyson, R.). New York: The Ronald Press Company, 1950.
11. Koffka, K. Perception: an introduction to the Gestalt-theorie. *Psychol. Bull.*, 1922, 19, 531-585.
12. Koffka, K. The growth of the mind. New York: Harcourt, Brace and Company, 1928, 108-111.
13. Koffka, K. Principles of gestalt psychology. New York: Harcourt, Brace and Company, 1935.
14. Köhler, W. The mentality of apes. New York: Harcourt, Brace and Company, 1927.
15. Köhler, W. Gestalt psychology. New York: Horace Liveright, 1929.
16. Köhler, W. Gestalt psychology: an introduction to new concepts in modern psychology. New York: Liveright Publishing Company, 1947.
17. Lewin, K. A dynamic theory of personality. New York: McGraw-Hill Book Company, Inc., 1935.
18. Taylor, H. A study of configurational learning. *J. comp. Psychol.*, 1932, 13, 19-26.

19. Titchener, E. B. A text-book of psychology. New York: The Macmillan Company, 1928, p. 37.
20. Watson, J. B. Behaviorism. New York: W. W. Norton and Company, 1925, p. 6.
21. Wheeler, R. H. The science of psychology. New York: Thomas Y. Crowell Company, 1929, p. 317.

5

THORNDIKE'S AND WOODWORTH'S EXPERIMENTS ON THE TRANSFER OF TRAINING AND THEIR INFLUENCE UPON THE DOCTRINE OF FORMAL DISCIPLINE

1. FORMAL DISCIPLINE AND FACULTY PSYCHOLOGY

THE DOCTRINE of formal discipline is the time-honored view that by hard study and application the "fiber of the mind" becomes toughened like a muscle; and that, as a direct result, one's "powers" of attention, memory, reasoning, and the like are markedly strengthened and increased. Advocates of this belief will contend that a boy's "reasoning ability," if thoroughly trained by the solving of numerous originals in geometry for example, is thereby better prepared to handle the knotty problems of business and professional life. In other words, the transfer of training from one activity to another, or the spread of improvement, is conceived to be broad and general. Opposed to the concept of general mind training inherent in the idea of formal discipline is the view that transfer of acquired skill or special training is relatively specific and narrow; that training, for example, in the duties of a department store salesman will carry over or benefit another activity or pursuit only insofar as there is a similarity of method or material between the two. For many years a controversy, which is still very much alive, has raged between these two opposing views.

Historically, the doctrine of formal discipline owes its origin and its present-day survival to two well-entrenched beliefs. The *first* is the view that mind is made up of numerous distinct powers or faculties—the so-called "faculty psychology." According to the extreme form of faculty psychology, mind is analogous to an intricate machine, the various parts of which correspond to the faculties of observation, memory, reasoning, judgment, volition, and the like. Experiences are the raw material fed into this mind-machine; here they are weighed and estimated by the faculty of judgment, arranged in logical sequence by the faculty of reasoning, and stored away by the faculty of memory. With training and long practice

the mind takes on strength, agility, and quickness, and these characteristics are generally exhibited in all sorts of situations.

The *second* belief holds that discipline is the essential function and even the main duty of education. This idea grows directly out of faculty psychology. It is deeply rooted, too, in the moralistic view that whatever is difficult—by virtue of that fact alone—is valuable in the training of the child's mind. This view is still widely prevalent today among school people, although it has shown some signs of weakening in recent years. It has long been a tradition that among school studies Latin and mathematics are especially useful in training the powers of concentration, reasoning, and precision. Growth of the belief in the disciplinary value of Latin is interesting to trace. For many years Latin was taught as a very necessary subject for priests and men of learning; later on it was studied for its cultural value in contrast to the often despised language of the common people. Today, when the chief use of Latin seems to be in furnishing inscriptions for monuments and in decorating college diplomas, its value has become largely disciplinary. Also, Latin's value in disciplining the mind has grown greater and greater as its utility has become less and less. Much the same thing is true of certain branches of mathematics. Thousands of high school boys and girls study plane geometry today, although an almost negligible percentage of them will ever use geometric principles in later life. Retention of geometry in the curriculum is based not upon its utility but upon the alleged value of pure mathematics in training the mind in precision, in judgment, and in logical reasoning.

The following quotations, some of which are quite recent, show how strong and orthodox is the belief in the superior disciplinary value of certain subjects. All of these statements reflect the view that an individual has a mind which possesses certain powers which can be trained and disciplined by hard and not always interesting tasks. It is apparent that this assumption is based implicitly, if not explicitly, upon faculty psychology.

The study of the Latin language itself does eminently discipline the faculties and secure to a greater degree than that of the other subjects we have discussed, the formation and growth of those mental qualities which are the best preparation for the business of life—whether that business is to consist in making fresh mental acquisitions or in directing the powers thus strengthened and matured, to professional and other pursuits.—*Lectures on Education*, by Joseph Payne, Vol. I, page 264.

The most valuable thing in the way of discipline which comes from the study of a foreign language is its influence in improving the pupil's command of his own. Of course this means the improvement in general judgment and discrimination which is evinced by a finer linguistic sense.—*Methods of Teaching the Modern Languages*, by E. H. Babbitt, page 126.

The value of the study of German "lies in the scientific study of the language itself, in the consequent training of the reasoning, of the powers of observation,



ROBERT S. WOODWORTH

1869-

comparison and synthesis; in short, in the upbuilding and strengthening of the scientific intellect."—*Methods of Teaching Modern Languages*, by Calvin Thomas, page 26.

We speak of the disciplinary studies . . . having in our thought the mathematics of arithmetic, elementary algebra, and geometry, the Latin-Greek texts and grammars, the elements of English, or of French or of German. . . . The mind takes fiber, facility, strength, adaptability, certainty of touch from handling them, when the teacher knows his art and their power. The college . . . should give . . . elasticity of faculty and breadth of vision, so that they shall have a surplus of mind to expend.—Woodrow Wilson, in *Science*, November 7, 1902.

A knowledge of a foreign language contributes in an unusual degree to the making of internationally minded, broad thinking, intellectually resourceful and contented citizens.—B. Q. Morgan, *The Place of Modern Foreign Languages in the American High School*, School and Society, Feb. 28, 1928.

Specifically, Latin and Greek become effective as educational instruments in at least seven ways:

By training in the essentials of scientific method: observation, comparison, generalization;

By making our own language intelligible and developing the power of expression;

By bringing the mind into contact with literature in elemental forms;

By giving insight into basic civilization;

By cultivating the constructive imagination;

By clarifying moral ideals, and stimulating right conduct;

By furnishing means of recreation.—F. W. Kelsey, *Latin and Greek in American Education*, New York: Macmillan, 1921, page 21.

Every great study is not only an end in itself, but also a means of creating and sustaining a lofty habit of mind; and this purpose should be kept always in view throughout the teaching and learning of mathematics.—Bertrand Russell, *Mysticism and Logic*, New York: Longmans, 1921, page 64.

I believe that education has been getting worse in this country. It is trying to spread out and teach too much—a smattering of everything. We will live to see the time when there will again be more emphasis on mental discipline, on toughening the fibers of the mind, on subjects that make the mind tackle a hard problem and work it through to an accurate conclusion.

If you go through high school and college taking easy courses you are not cheating anyone but yourself. Put into your curriculum some hard courses—some courses that may have no relation to the work that you will do after but will exercise your mind, so that you can stand competition in the office or profession where you will eventually have to battle for a living.—Bruce Barton, 16th Annual Convention of CSPA, Saturday, March 16, 1940.

2. EVIDENCE AGAINST FACULTY PSYCHOLOGY

The modern view of mental organization holds, in opposition to the faculty doctrine, that what we loosely call "mind" consists of numerous more or less closely connected behavior tendencies, and that transfer of

improvement as a result of special training takes place almost exclusively among correlated activities. Activities may be related through similarities in form or in procedure and method, or through identities in content or material. A knowledge of Latin will undoubtedly aid one to learn Spanish more easily, while a knowledge of mathematics is indispensable in engineering or physics. Moreover, experience gained as a consulting chemist with one concern will usually enable a man the more easily to perform his professional duties with another firm. Added to this fairly specific transfer, there is almost certainly a general carry-over growing out of such broad factors as a common language, definite rules of study, habits of neatness, care, and thoroughness, quickness in reading and writing, and the like. Attitudes, motives and drives may also be general to a considerable degree.

The evidence against formal discipline and the faculty psychology comes from everyday observation, from physiology, and chiefly from experiment. We all believe in a certain amount of transfer—but in a certain amount only. The able lawyer is made a judge, and the active young executive becomes a general manager. But the engineer is not made health officer, nor is a minister or physician put in charge of building bridges. In other words, common sense and everyday observation tell us that special training cannot possibly prepare a man for all sorts of vocations nor make him an expert in all branches of knowledge. In fact, striking instances of the opposite are found. All too often the eminent scientist gives opinions on theology which are simply banal, while the highly trained mathematician is rarely more logical or rational in selecting an automobile than is his less gifted neighbor. In short, there is little evidence from everyday life of general mind training.

The pseudo-science of phrenology is closely related to faculty psychology on the physiological side. Phrenology taught that the different faculties of the mind have "seats" in specific parts of the brain, and that their respective developments can be estimated from the "bumps" or prominences on the skull (p. 177). Physiologists have found no evidence of nerve centers corresponding to the alleged phrenological faculties. They have been able to map out on the cortex general sensory and motor areas, a visual area, an auditory area, and an area for taste and smell. But no centers for such faculties as judgment, observation, volition, or concentration have ever been found. To the present-day psychologist the term *concentration*, for instance, instead of meaning a power or faculty, is simply a name to describe responses which are alike in being marked by a high degree of intensity. Such responses may fall into many categories, from concentration upon a symphony to concentration upon a boxing match. What a man "concentrates" upon is dependent upon his

age, intelligence, education, special interests, and training, not upon a mysterious "power of concentration" which exists *per se*.

The experimental attack upon the question of general *versus* specific transfer of improvement has been made by psychologists interested in the theoretical problem of mental organization as well as by those interested in the educational aspects of the transfer problem. As we shall see in the following sections, laboratory studies of transfer have thrown light upon both of these questions.

3. FIRST EXPERIMENTAL ATTACKS UPON THE PROBLEM OF TRANSFER OF TRAINING

The first experiment dealing with the problem of transfer seems to have been made by William James in about 1890. James measured the time required by himself and four students to memorize selections from the poems of Victor Hugo and Tennyson before and after training with other passages of poetry, for example, lines from *Paradise Lost*. After memorizing a given passage, each subject spent a varying time up to a month attempting to train his memory with other selections, after which he again memorized passages from the original material. On these final tests, three of the subjects showed slight improvement over their first records, but the other two were slightly worse than before training. This experiment of James was inadequate in method and inconclusive in result, but it possesses historical significance, since it started the experimental attack upon the problem of the extent of transfer through special practice. A later and better controlled series of experiments upon the general topic of memory training was carried out by W. F. Dearborn. Dearborn set out to discover whether the memorizing of lists of nonsense syllables and of meaningful material (e.g., vocabulary lists and poetry) would improve one's memory for numbers, letters, words, prose and poetry. The extent of transfer—i.e., the improvement in memorizing *after* practice—varied from zero to about 50 per cent. This degree of transfer was not so great as the improvement shown in learning the material upon which the subjects *specifically* practiced. Dearborn writes: "The results indicate that a considerable part of the improvement found must be attributed to direct practice in the test series, and not to any 'spread' of improvement from the practice series proper."

The most significant of the early attacks on the problem of transfer was a series of three studies by E. L. Thorndike and R. S. Woodworth. These experiments furnished the first exact and quantitative evidence against the doctrine of formal discipline. In order to get material which was not too familiar, tasks were drawn from the field of discrimination and

perception. The influence of the training received in *one* kind of perception upon efficiency in *other* related activities was then investigated by giving training to five or six subjects in (1) estimating the areas of various geometric figures, the lengths of lines, value of weights, and so forth; and (2) observing designated words or letters on a printed page.

In the first series of experiments—those on sensory and perceptual discrimination—Thorndike and Woodworth trained their subjects to estimate the areas of rectangles and other geometrical figures. *Before* the training or practice period, and again *afterward*, the subjects were tested for their accuracy in estimating the areas of other figures, e.g., rectangles, circles, triangles, and trapezoids. Some of these figures were different in shape but similar in size to those of the training period, others similar in shape but different in size, and still others different in both size and shape. From the records before and after practice, the effect of the intervening training period upon the accuracy of discrimination was determined. If we let *IT* = initial test; *TP* = training period; and *FT* = the final test, we can represent the experiment diagrammatically as follows:

IT—TP—FT

During the training period (*TP*) of the principal discrimination experiment, the subjects practised in estimating the areas of rectangles from 10 to 100 square centimeters in size, in every case carefully checking the accuracy of their results until they had acquired marked proficiency at this task. The improvement—as measured by tests given before and after this training period—made in estimating areas of the *same size* as these rectangles, but of *different shape*, was not quite one half as large (44 per cent) as the improvement shown in the *TP*, namely, in estimating areas *different* in size but of the *same shape*. When subjects were tested with figures of the *same shape* but somewhat *larger* in size (e.g., rectangles of 140 to 300 square centimeters) than those used in the training series, the improvement was only one third as great (30 per cent) as in the *TP*. Very probably this low figure is partly a result of confusion and interference due to the similarity in shape of the rectangles in the two series. For areas different in shape and larger in size (140 to 400 square centimeters), the improvement in accuracy was 52 per cent as great as in the training series.

Thorndike and Woodworth next trained their subjects in estimating the lengths of lines .5 to 1.5 inches long until an improvement of about 25 per cent was noted. This training period produced little or no improvement in the estimation of lines six to twelve inches long as shown by tests given before and after the practice period. Nor did the training with six-inch to twelve-inch lines lead to any improvement in estimating lines fifteen to twenty-four inches long.

In another sensory field, that of tactual and kinesthetic discrimination, improvement to about the same degree as with geometrical figures was noted as a result of previous training in closely similar functions. The subjects were tested this time for their accuracy in estimating weights of 120 to 1,800 grams. They were then trained in estimating weights from 20 to 120 grams; and retested with the first set of weights in order to measure the effect of the intervening practice period. The improvement in the *FT* over the *IT* was about 40 per cent of that improvement which appeared in the *TP* alone, indicating that transfer, while evident, was by no means considerable.

Thorndike's and Woodworth's second series of experiments, those with language, are especially valuable because they parallel fairly closely the kind of work done in school from which much transfer has been confidently expected. In these experiments, which dealt with the perception of words and letters, the training period consisted in the rapid cancellation of every word on a printed page containing both *e* and *s*. *Before* and *after* this training, the subjects were tested for their ability to perceive and cancel words containing *a* and *t*, *s* and *p*, *l* and *o*, and the like. In another experiment the subjects were tested for their ability to perceive and cancel certain parts of speech, e.g., verbs and adjectives, as well as words of a certain length, and misspelled words. They were then trained in like material and again retested. Improvement in finding and canceling two letters—as shown by tests given before and after training—was about 25 per cent as great as the accuracy achieved in the training period itself. Efficiency in finding and marking verbs and prepositions was improved from 20 to 25 per cent in speed, and often three times as much in accuracy, as a result of training with analogous material. But although the ability to perceive and mark other parts of speech as a result of this training in some instances showed an increase in speed, there was often a great decrease in accuracy. Here again interference rather than transfer of improvement seems to have been the rule.

The result of Thorndike's and Woodworth's experiments indicated that even in performances superficially alike there may often be not only comparatively little transfer of improvement, but instead considerable interference or negative transfer. Such transfer as appeared our authors attributed to the carry-over of specific methods and rules of procedure, or to similarity in the trained and tested material. One subject, for example, discovered in himself a tendency to overestimate small areas, and allowed for this in his later judgments. Another found that too much interest in the subject-matter of the selection canceled slowed him up, and that too close attention to the particular letters canceled led to interference when other letters were marked. In general, attitudes of confidence,

familiarity with the task, and improvement in methods of attack seem also to have been transferred.

Thorndike's and Woodworth's experiments delivered a broadside against mental discipline—at least in its extreme form—from which it never fully recovered. These experiments gave little encouragement to the belief in a general observational ability or to confidence in a faculty of perception which, once it has been trained, leads thereafter to a general increase in the level of performance in all perceptual activities. On the contrary, these experiments suggested strongly that the *amount* of transfer from one activity to another depends mainly upon the degree of community between the two tasks. The dozens of experiments on transfer carried out since—although differing in method of attack and in activity selected for study—have on the whole corroborated Thorndike's and Woodworth's findings. The transfer effect of practice has proved to be far from general and is, for the most part, confined to related capacities. Such a result is, of course, diametrically opposed to the idea of formal discipline.

4. TRANSFER EXPERIMENTS IN MAZE LEARNING

The greatest amount of transfer seems to be found in studies of animal maze learning, the least in studies of the disciplinary or transfer value of school subjects. Typical investigations from each of these fields will be cited in this and the next sections. A careful and well controlled study of transfer in maze learning (in which both rats and humans were subjects) was carried out by Webb at the University of Chicago. Webb employed six square mazes as learning problems for his animals. In order definitely to favor positive transfer, two of these mazes, *A* and *B*, were closely similar as to turns and blind alleys, while the other four, *C*, *D*, *E*, and *F*, were dissimilar in this respect in order to favor negative transfer. When a group of rats had learned maze *A*, transfer was measured by seeing how much more quickly these trained rats could learn the other mazes than comparable control groups which had had no previous practice in *A*. Records were kept in terms of number of trials, errors made, and time required to learn. In every instance, there was positive transfer from maze *A* to the other five mazes, the greatest carry-over being from maze *A* to maze *B* wherein the patterns were *most similar*. The effect of practice was shown in a tendency to enter blind alleys less frequently. Increase in the efficiency of learning the other mazes as a result of practice in *A* varied from 19 to 77 per cent for the trials required; from 20 to 95 per cent for errors; and from 29 to 90 per cent for time. The greatest transfer occurred in the first five trials with the second maze. Substantially the same result, i.e., consistently positive transfer, was found for human subjects, who were trained with four pencil mazes.

Positive results in this experiment are probably to be attributed to the fact that mazes require the learning of fairly simple motor habits. Even the most dissimilar mazes are more alike than are the learning tasks ordinarily confronting rats or human beings outside of the laboratory. The fact that transfer of practice is concentrated mainly in the first five trials on the second maze is especially significant. It suggests that what is carried over from maze A to the other mazes is a *general familiarity* with the task to be learned, plus feelings of confidence, perhaps, which enable the learner (rat or human being) to get off to a "running start." Transfer even from maze A to maze B, however, fell considerably below 100 per cent, the average carry-over being about 82 per cent.

The positive transfer found in this experiment was corroborated by Wiltbank who repeated and extended Webb's work, using different mazes. An interesting result reported by Wiltbank is that there is little transfer *except* to a maze which is unfamiliar or only slightly learned. Different groups of rats were given 0, 2, 4, 8 or 16 trials in maze 1. They then learned maze 2 *completely*, after which maze 1 was completely learned. Table II shows that when the first maze is unknown or slightly known (0 to 2 trials) there is a large saving in trials, errors, and time as a result of learning a second maze. When the first maze is fairly well known (8 to 16 trials), there is little transfer from a second maze:—in fact, negative transfer or *interference* is more likely than positive transfer. The appearance of negative transfer means that the carry-over even from one simple motor habit to another may prove to be a hindrance rather than a help. The problem, then, is not simply one of transfer *versus* no transfer; but of transfer *versus* no transfer plus interference!

TABLE II

Number of Trials in 1st Maze before Learning 2nd Maze	Saving in Total		
	Trials	Errors	Time
0	43	61	53
2	45	55	19
4	33	-5	-8
8	22	-33	-32
16	-11	-43	24

5. TRANSFER EXPERIMENTS IN SCHOOL LEARNING

Probably because the advocates of formal discipline have so vociferously defended the training value of Latin, many experimenters have appeared to be especially eager to test the claims for this subject. Several studies have been made of the transfer effect of Latin. In one of these the gains in English vocabulary during the first year of high school, made by 717

Latin pupils, were compared with the gains in vocabulary made by 677 non-Latin pupils. Two vocabulary tests were administered, one at the beginning and the other at the end of the school year; the first consisted of twenty-five English words derived from Latin words frequently occurring in high school Latin, and the second of twenty-five words derived from Anglo-Saxon and Greek sources. The gains made by the pupils upon the words of Latin and non-Latin origin are shown in Table III.

TABLE III

VOCABULARY GAINS DURING FIRST YEAR OF HIGH SCHOOL OF 717 LATIN PUPILS AND 677 NON-LATIN PUPILS UPON (A) 25 WORDS DERIVED FROM LATIN AND (B) 25 WORDS DERIVED FROM ANGLO-SAXON AND GREEK SOURCES

	<i>Latin Derivatives</i>	<i>Non-Latin Derivatives</i>
717 Latin pupils	5.5	1.2
677 Non-Latin pupils	2.0	1.4
	<u>3.5</u>	<u>-0.2</u>

The Latin pupils improved more than did the non-Latins in their knowledge of English words of Latin origin; but on the non-Latin words neither group showed significant improvement. The study of Latin evidently made the student better acquainted with certain English words derived from Latin. But apparently it did not develop any "general language ability" which operated to improve the students' vocabulary for words of non-Latin origin.

It is interesting to see to what extent method of teaching affects the transfer value of a school subject. In one experiment the gains in English vocabulary made by four groups were compared. The four groups included a class in English conventionally taught; a class in English taught with emphasis upon word study; a class in Latin conventionally taught, and a class in Latin taught with special reference to word meanings and derivations. The children in the four groups were approximately equated for general intelligence by means of mental tests. Gains in English vocabulary made by the various classes are shown in Table IV:

TABLE IV

COMPARATIVE GAINS IN ENGLISH VOCABULARY OVER ONE YEAR MADE BY FOUR CLASSES DIFFERENTLY TAUGHT

	<i>Gain</i>
Latin with word study	10.7
English with word study	8.5
Latin, conventionally taught	6.7
English, conventionally taught	6.0

When attention was *specifically* directed toward derivatives and word meanings, Latin "transferred" to English vocabulary—being slightly more effective than English. When there was no especial emphasis upon word study, there was little difference between Latin and English, neither being very effective.

The most extensive study of the transfer value of various school subjects was carried out by Thorndike in 1924. This experiment meets the criticism often leveled (correctly) against laboratory experiments on transfer (1) that training is too brief and (2) that the functions studied are too narrow for comparison to be made with the training given in school or life. Thorndike investigated the effect upon intellectual achievement of a year's work in various high school subjects, such as Latin, mathematics, and history. The subjects were 8,564 high school students in grades 9, 10, and 11. As a preliminary or initial test (*IT*) all pupils were given Form A of the *Tests of Selective and Relational Thinking*, published by the Institute of Educational Research, Teachers College, Columbia University.¹ After one year of school work these students were given Form B of the same examination as a final test (*FT*). The two forms A and B are equivalent, being of approximately equal difficulty. The training period (*TP*) was supplied by one regular year of schooling in from four to five studies. The gain in final score over initial score (*FT-IT*) was taken to be a joint product of growth *plus* the improvement or transfer effect of the subjects studied. To separate out the differential effect of each school subject, Thorndike employed the following plan. All pupils whose programs included, let us say, English, geometry, history, and Latin were matched for initial score in Form A with pupils whose programs included English, geometry, history, and some other subject than Latin, say physics. These two groups, being equal in initial ability, and having taken the same subjects throughout the year, except for Latin in one group and physics in the other, any difference in gain upon Form B at the end of the year would be the result, presumably, of the differential or transfer effect of the two compared subjects. To cite an example, if the average gain in the final test of the Latin Group is 25 points, and the average gain of the Physics Group 15 points, then since the two groups were equal at the start, the training effect from Latin is, on the average, ten points greater than that from physics. The transfer effect of the other school studies was calculated in the same way. The amount which each contributed to the final score was computed by balancing programs alike except for the one variable study in each case. Of course, various com-

¹ This "battery" included those tests usually found in group tests of general intelligence, e.g., arithmetic, opposites, number series completion, analogies, sentence completion, etc. There were fifteen tests in all.

binations of studies had to be made owing to the wide diversity of programs, so that the task of determining differential gains was by no means so simple as our illustration might suggest. The example given will serve, however, as a much simplified description of the method. To determine the carry-over value of a study as against simple growth effects during the year, equated groups with programs containing studies *ABCD*, for instance, and studies *ABCDX* were compared. This gives the effect of study *X* against non-*X*, the growth period being the same for both groups of students.

The results of this extensive investigation were later checked in another study of 5,000 pupils. The implications of these two experiments for formal discipline are extremely important, and the specific findings are probably at variance with what many of us might have been led to expect. Table V summarizes the results for more than 13,000 pupils. As far as the different studies are concerned, mathematics proved to have the greatest training effect, with the social sciences, civics, economics and the like a close second. Latin is inferior to mathematics and science, about equal to French, and superior to sewing, stenography, manual training, and dramatic art. In general, these last-named studies show a slight negative transfer—a *loss* in final score rather than a gain. Clearly the traditional view that Latin is the subject par excellence for training one to reason or think is not borne out by these findings.

TABLE V

RELATIVE VALUE OF A YEAR'S TRAINING IN VARIOUS HIGH SCHOOL SUBJECTS IN INCREASING "ABILITY TO THINK"

<i>Subject Groups</i>	<i>Weighted Average Differences (Relative Influence)</i>
1. Algebra, geometry, trigonometry, etc.	2.99
2. Civics, economics, psychology, sociology	2.89
3. Physics, chemistry, general science	2.71
4. Arithmetic and bookkeeping	2.60
5. Physical training	.83
6. Latin, French	.79
7. Cooking, sewing, stenography	— .14
8. Biology, zoölogy, botany	— .15
9. Dramatic art	— .48

Fairly consistent differences appeared as between school subjects, but the transfer effect of even the best subjects was astonishingly small. The gain made by a pupil during the year apparently depends far more upon his native ability (as shown by a high initial score on Form A) than upon the transfer value of a particular study. This is clearly shown by the fact that the *highest* 1 per cent in initial general ability gained about twenty

points in the final test after a year's work in high school, while the *lowest* 1 per cent gained only one and one-half points. This gain was irrespective of the subjects studied. In commenting upon these results Thorndike writes:

The expectation of any large differences in general improvement of the mind from one study rather than another seems doomed to disappointment. The chief reason why good thinkers seem superficially to have been made such by having taken certain school studies, is that good thinkers have taken such studies, becoming better by the inherent tendency of the good to gain more than the poor from any study. When the good thinkers studied Greek and Latin, these studies seemed to make good thinking. Now that the good thinkers study Physics and Trigonometry, these seem to make good thinkers. If the abler pupils should all study Physical Education and Dramatic Art, these subjects would seem to make good thinkers. These were, indeed, a large fraction of the program of studies for the best thinkers the world has produced, the Athenian Greeks. After positive correlation of gain with initial ability is allowed for, the balance in favor of any study is certainly not large. Disciplinary values may be real and deserve weight in the curriculum, but the weights should be reasonable.

Thorndike's findings uncovered a common fallacy into which many staunch defenders of the disciplinary subjects (especially the defenders of Latin and mathematics) fall. This is the fallacy of *selection*. Because more classical students than non-classical go to college, take honors, make Phi Beta Kappa, and later appear in *Who's Who*, superior achievement is taken to be a *result* of classical training. The fact is overlooked that it is the more intelligent students who take Latin and mathematics. It is likely that this intellectual superiority leads to later achievement rather than the classical training itself. Given high intelligence to begin with, almost any kind of training is effective, as Thorndike's results showed.

6. THEORIES OF TRANSFER: SUMMARY AND CONCLUSIONS

The present-day opinion of psychologists on the problem of transfer of training may be summarized as follows: improvement in one function, as a result of the exercise of another, depends (1) upon the identity of elements or components (to be found both in the material and method) between the activities being measured; or (2) upon the carry-over of attitudes or methods or techniques of attack which are generalized. Both of these descriptions of the mechanism of transfer, it need hardly be said, are directly opposed to the faculty notion of mental organization. The first view is substantially that of Thorndike and Woodworth stated above; the second, that of "generalized experience," was first advanced by Judd and is sometimes taken to be an alternative—and opposed—explanation of transfer. It is doubtful, however, whether there is any fundamental

opposition between the two theories, though there is a difference in emphasis. The value of Latin as an aid in learning French depends upon the "identical elements" in the two languages, as well as upon the similarity in general form and syntax which they possess in common with English. Common attitudes and techniques are best abstracted from a variety of experiences when there are identical bonds running through them all. The most common bond is language; others are everyday information, similarity of method or procedure, and plan of attack. Such connections as these supply the identical elements as well as the more abstract principles upon which transfer depends.

Some psychologists have objected to the notion of transfer through identical elements. They hold that such a view overemphasizes the mechanical linkages among overlapping habits, and does not take sufficient account of the intelligent coöperation of the individual, his interests, motivation, and drives. Allport comments, apropos the theory of identical elements, that it is perhaps only the young child and the mentally undeveloped who must be taught in terms of specific habits because not capable of generalizing. Part of the trouble here lies in the use of the term "element." Undoubtedly, any theory which deals with elements, thought of literally and perhaps atomistically as numerous and shifting bits of behavior (like grains of sand), is much too narrow to cover a great many cases of transfer. The man with a "chip on his shoulder" sees insults in many apparently innocuous situations; the religious fanatic interprets even disasters in pious phrases; the "one hundred per cent American" sees any opposition to his views as evidence of subversive activity; the incurable optimist always sees the "silver lining" no matter how tarnished it may be. In each of these illustrations, it is evident that the "identical components" or linkages are supplied by the emotionally colored glasses of the subject—though there may be other common elements as well. Besides a common motivation and similar emotional interests, superior intelligence makes for transfer (as we have seen, p. 94) by supplying relations and analogies as well as general principles of attack. In short, the identical elements or common bonds may be supplied by the behaving person as well as by the situation. Viewed thus broadly, there is no contradiction between a theory of identical elements and a view which sees transfer as dependent upon the personality organization of the individual.

A summary of studies of transfer of training is presented in Table VI. This table based upon the results of many investigations justifies the conclusions which follow.

- (1) *Some* transfer of training is the rule rather than the exception.
- (2) The degree of transfer depends upon the *amount* and *kind* of training; teaching specifically for transfer enhances the carry-over effect.

- (3) Transfer depends upon native intelligence and upon the development of techniques and methods.
- (4) Transfer may depend upon a unifying motive or drive.
- (5) While some school subjects have greater transfer value than others, the transfer effect of none is very large.

TABLE VI
SUMMARY OF TRANSFER EXPERIMENTS
(After Orata)

Finding	Number of Investigations			Per Cent
	Laboratory Experiments	Classroom Experiments	Total	
Finds considerable transfer	18	14	32	32
Finds appreciable transfer	23	26	49	49
Finds very little transfer	1	7	8	8
Finds no transfer	1	1	2	2
Claims transfer, but no data given . .	1	0	1	1
Claims no transfer, but no data given	1	1	2	2
Claims no transfer, but faulty calculation	0	2	2	2
Finds transference and interference	8	0	8	8
Finds interference only	3	0	3	3
Grand Total	56	51	107	108
Duplication	8	0	8	8
Net Total	48	51	99	100

7. THE EXPERIMENTAL TECHNIQUE IN TRANSFER EXPERIMENTS

Many of the early experiments on the transfer value of practice were faulty because of the experimenter's failure to check the results obtained from special training against the results secured from a *control group*. A control group, which takes the initial test and the final test, but not the training, enables one to separate out, from the transfer effect due to special training, the carry-over due to mere repetition of the final test. The control group technique—as illustrated by the *method of equivalent groups*—is often employed in transfer experiments as well as in many other experimental problems in psychology. The first step is to match the two (or more) groups on the basis of one or more variable factors, such as, for example, age, sex, or general intelligence. Often the groups are equated in terms of the specific function, e.g., memory or reaction time, upon which the influence of the *experimental factor* or *EF* (in transfer experiments this is special practice) is to be studied. In equating groups, the usual method is to pair off the subjects so that the two groups will

have the same or nearly the same average initial score *IT* and the same variability around this mean. When the groups have been equated, the *EF* is applied to one group—the experimental—and at the conclusion of this training period *both* groups are given the final test *FT*. The difference in gain of the experimental group over the control gives a measure of the effect of the *EF*, *minus* the practice effect of simply repeating the test. That is to say, the control group enables us to sift out the practice effect (due to mere repetition of the *IT*) from the effect of the *EF plus* the repetition of the *IT*. We may diagram the whole process as follows:

$$\begin{aligned}\text{Experimental group: } IT - EF - FT &= C_1 \\ \text{Control group: } IT - \text{———} - FT &= C_2 \\ C_1 - C_2 &= \text{the effect of the } EF\end{aligned}$$

An experiment by Gates and Taylor illustrates clearly the method of equivalent groups as applied to transfer problems; it is cited here as a concrete example of procedure. The problem was to discover to what extent a simple function, i.e., digit span, could be improved as the result of intensive practice and how permanent such improvement would be. A group of kindergarten children four to six years old were subjects. From this group two equated groups were made up by matching each child with another child as nearly as possible in the following traits: sex, age, mental age, IQ, scholastic maturity as estimated by teachers, memory for digits presented orally, for letters presented orally, for unrelated words presented orally, for geometric figures, for pictures, and for names. How well the two groups were matched may be judged from the following table:

TABLE VII

	Number	Age	Mental Age	IQ	Memory for Digits	Memory for Letters	Memory for Unrelated Words	Memory for Related Words	Memory for Geometrical Figures	Memory for Pictures	Memory for Names
Experimental group	16	5.1	6.31	122	4.33	3.64	3.86	14.0	4.3	5.3	7.5
Control group	16	5.1	6.35	123	4.33	3.71	4.07	13.7	4.0	5.7	7.0

The experimental factor was intensive practice in immediate memory for digits. For seventy-eight days the children in the experimental group were given daily drill in learning groups of digits, the control group, meanwhile, receiving no training. At the end of the practice period when both groups were retested (*FT*) for digit span, the practice group had progressed from an average span of 4.33 digits (see Table VII) to an average of 6.40, a gain of 2.07 digits. Meanwhile the control group had

only increased its average from 4.33 to 5.06, a gain of .73 digit. The difference between 2.07 and .73 gives the effect of the intervening practice upon this simple memory function. Gates writes: "In the Stanford-Binet Scale, repeating 4 digits is placed at year 4, and 6 at year 10. The practice group, then, advanced during a period of 4.5 months during which they practiced on 78 days, an amount equal to that which the average untrained child advances in approximately 6 years."

Does this comparatively enormous gain result from real improvement in retentivity, or from a stimulated rate of growth of neural connections, or can it be attributed to an improvement in the technique of memorizing, better methods and the like? To answer this question, Gates allowed four and a half months to pass, and then retested fourteen members of each group—all that could be located. The result was quite conclusive. On this retest, the average memory span of the practiced group was 4.71 digits, that of the control group 4.77 digits: the two groups were equal again as they were at the beginning of the study. Improvement, although quite large immediately after training, had resulted in no permanent benefit. This result indicates clearly that the large increase in score brought about by the seventy-eight days of intensive practice was probably due almost entirely to the acquisition of special techniques, familiarity with the task in hand, adaptation to the examiner's voice and signals, loss of anxiety, better habits of attention, and the devising of simple schemes of grouping, forming combinations, and the like. These techniques, though temporarily highly effective, did not remain as permanent acquisitions.

The method of equivalent groups has proved to be exceedingly valuable in studying the effects of various experimental factors. Apart from the *EF* of special training in transfer experiments, the incentive value of praise and blame, the effects of various drugs upon mental activity, visual *versus* auditory presentation, and positive *versus* negative suggestion may be mentioned as a few of the many variables which have been investigated by this method.

BIBLIOGRAPHY

1. Allport, G. W. *Personality: a psychological interpretation*. New York: Henry Holt and Company, 1937.
2. Brolyer, C. R., Thorndike, E. L., & Woodyard, E. A second study of mental discipline in high school studies. *J. educ. Psychol.*, 1927, 18, 377-404.
3. Davis, R. A. *Psychology of learning*. New York: McGraw-Hill Book Company, Inc., 1935, chapter 10.
4. Dearborn, W. F. Experiments in learning. *J. exp. Psychol.*, 1910, 1, 373-388.
5. Gates, A. I., & Taylor, G. A. An experimental study of the nature of improvement resulting from practice in a mental function. *J. educ. Psychol.*, 1925, 16, 583-592.

6. James, W. The principles of psychology. New York: Henry Holt and Company, 1890. Two volumes.
7. Judd, C. H. The relation of special training and general intelligence. *Educ. Rev.*, 1908, 36, 28-42.
8. Morgan, B. Q. The place of modern foreign languages in the American high school. *Sch. & Soc.*, 1928, 27, 185-193.
9. Orata, P. T. The theory of identical elements. Columbus, Ohio: Ohio State University Press, 1928.
10. Pressey, S. L. Psychology and the new education. New York: Harper & Brothers, 1933, chapter 15.
11. Starch, D. Educational psychology. New York: The Macmillan Company, 1919, 201-202.
12. Symonds, P. M., & Penney, E. M. The increasing of English vocabulary in the English class. *J. educ. Res.*, 1927, 15, 93-103.
13. Thorndike, E. L. Mental discipline in high school studies. *J. educ. Psychol.*, 1924, 15, 1-22, 83-98.
14. Thorndike, E. L., & Ruger, G. J. The effect of first year Latin upon knowledge of English words of Latin derivation. *Sch. & Soc.*, 1923, 18, 260-270, 417-418.
15. Thorndike, E. L., & Woodworth, R. S. The influence of improvement in one mental function upon efficiency of other functions. *Psychol. Rev.*, 1901, 8, 247-261, 384-395, 553-564.
16. Webb, L. W. Transfer of training. *Psychol. Monog.*, 1917, 24, No. 3.
17. Wilson, Woodrow. Princeton for the nation's service. *Science*, 1902, 16, 721-731.
18. Wiltbank, R. T. Transfer of training in white rats upon various series of mazes. *Behav. Monog.*, 1919, 4, No. 1.
19. Woodworth, R. S. Experimental psychology. New York: Henry Holt and Company, 1938.

EBBINGHAUS'S STUDIES IN MEMORY AND FORGETTING

1. EBBINGHAUS AND THE INVENTION OF NONSENSE SYLLABLES

THE IMPORTANCE to experimental psychology of the memory experiments of Hermann Ebbinghaus lies in the intrinsic value of the work itself as well as in the impetus and inspiration which it gave to literally scores of later investigations. Ebbinghaus was born in 1850 near Bonn, Germany. He attended several German universities, but returned to take his doctor's degree in philosophy at the University of Bonn at the age of twenty-three. For eight years (1886-1894), Ebbinghaus was a professor at the University of Berlin, from which post he went to the University of Breslau. He died in 1909 while in the midst of preparing a third edition of his book *Principles of Psychology*.

In his experimental work on memory, Ebbinghaus was much influenced by Fechner,¹ from whom, apparently, he got the idea of mental measurement. The results of his memory experiments were published in 1885 after several years of interrupted work. They mark the first real attempt to apply precise scientific method to the study of the "higher mental processes." Such phenomena had been regarded as too "subjective," too fleeting, even too personal, perhaps, for exact and quantitative treatment. It is Ebbinghaus's contribution to have shown conclusively that memory products are as amenable to experiment and to measurement as are any of the other natural facts with which science deals.

Ebbinghaus devised several methods for measuring memory which can be employed with different kinds of material. In addition, he introduced a new kind of memory device—the so-called *nonsense syllables*—which possesses among other advantages the very real one of being relatively free from "ready-made" associations. Nonsense syllables as constructed by Ebbinghaus were meaningless combinations of three letters, each combination consisting of two consonants separated by a vowel or diphthong. Examples are *bap*, *tox*, *muk*, and *rif*. Nonsense syllables of

¹ For Fechner's work on the measurement of sensation, see Chapter 16.

four letters, e.g., *nult*, *rulb*, *selx*, are perhaps more widely employed in memory experiments at the present time than are those of three letters.

The very real advantage of using nonsense syllables in memory experiments appears when one sets out to measure the memory ability of a small group of educated adults, using as material Lincoln's *Gettysburg Address*, say, or Poe's *Raven*. Differences in memory ability will quickly appear, a considerable part of which may be attributed to the degree of acquaintance with the material possessed by the subjects. In fact, probably as much individual variation will result from differences in familiarity as from differences in native ability to learn. One man may be favored because he is already familiar with the selection, or because he is of literary bent and hence accustomed to such memory tasks. Another man may be at a decided disadvantage because he is a mechanical engineer in whose everyday activities literary selections and poems are rarely encountered. This unequal initial state of acquaintanceship will always exist for prose materials or for poetry when subjects differ in age, education, cultural background, or training; and it is present, although to a somewhat lesser extent, even with children.

It was to overcome these handicaps, and as far as possible to equalize the backgrounds of different learners so that each might start with a clean "memory slate," that Ebbinghaus invented his nonsense syllables. An equally important advantage of nonsense syllables is that the same individual can be tested under different conditions and at different times with material sufficiently equal in memory value and homogeneous in content as to yield closely comparable results. All told, Ebbinghaus constructed about 2,300 nonsense syllables which he employed in the experiments to be described.

2. EBBINGHAUS'S EXPERIMENTS ON MEMORY

The phenomena of memory may be conveniently classified under the heads of (1) fixation, (2) retention, (3) recall, and (4) recognition. Fixation refers to the acquisition process; it is concerned with getting the impression, learning the new activity. Retention is usually thought of in physiological terms (brain traces, neurograms, etc.); it can be measured only indirectly through recall and recognition. Both recall and recognition measure permanence of learning. The distinction between them will appear in the following example. One recalls facts in answer to a general question ("When was the Battle of Waterloo?"); he recognizes the "right" answer as one response out of several possible responses ("Which of these five pictures have you seen before?").



Bettmann Archive

HERMANN EBBINGHAUS

1850-1909

Psychologists have been chiefly concerned with methods of insuring fixation, as responses must first be learned or established before they can be recalled or recognized. Since Ebbinghaus's pioneer work, methods of presentation, techniques of organizing material, and the relation of fixation to recall have been carefully investigated. Ebbinghaus was primarily interested in the problems of fixation and recall. All of his experiments were carried out upon himself as subject and were performed in an extremely careful and highly controlled manner. In each group of studies, Ebbinghaus proposed a general question which he attempted to answer by means of specific experiments. These problems may be arranged under the following five heads:

1. What is the relationship between the amount of material to be memorized and the time and effort required to learn it? Specifically, what effect does the *length* of a series of nonsense syllables have upon the *rapidity* with which it can be memorized?

2. What is the relationship between amount or degree of learning and retention? What effect does the *number* of repetitions of a given series have upon its retention?

3. How is *forgetting* related to the time-interval between learning and recall? What effect does the passing of time have upon one's memory for comparable series of nonsense syllables?

4. What effect do *repeated* learning and frequent review have upon one's ability to retain what he has studied?

5. What sort of connections are formed in learning: do they run forward serially from one term to the next following only, or do they skip terms in the forward direction, and even sometimes in the backward direction? What is the relative strength of these different associations, assuming them to be so formed?

These problems may appear at first glance to be pretty closely restricted to nonsense-syllable learning, and hence to be of minor value in clearing up the vexed questions of learning and forgetting so common in everyday life. Closer inspection of Ebbinghaus's work, however, will show that its bearing upon complex problems of memory is direct and considerable. Ebbinghaus approached such topics in the only truly scientific way, namely, through the use of a rigorously controlled method and of standard material which avoided to a high degree inequalities in initial memory value.

In the following sections we shall consider the five experimental problems, listed above, and show how they were attacked by Ebbinghaus.

3. THE AMOUNT OF MATERIAL TO BE MEMORIZED AND THE TIME AND EFFORT REQUIRED TO LEARN IT

Common experience tells us that the longer a poem or a prose memory lesson, the harder it is to learn up to the point where it can be repeated "by heart." Will the learning of ten verses of a poem require twice as long, three times as long, or six times as long as the learning of five verses? Probably no one would care to commit himself to a definite arithmetical answer to such a question as this, though doubtless every one would be perfectly certain that the longer selection would take more time. Ebbinghaus attacked this problem in the following way. First, he recorded the time and the number of repetitions necessary for him to learn different lists of seven, ten, twelve, sixteen, twenty-four, and thirty-six nonsense syllables each, up to the point of *one* errorless reproduction. This scheme Ebbinghaus called the "learning method." He then calculated the average time spent on each syllable in the different lists, in order to get a comparative measure of effort. Ebbinghaus's results may be stated in tabular form as follows:

TABLE VIII

<i>Length of Lists</i>	<i>Number of Readings</i>	<i>Time for Lists</i>	<i>Average Time per Syllable</i>
7	1	3 secs.	.4 secs.
10	13	52	5.2
12	17	82	6.8
16	30	196	12.0
24	44	422	17.6
36	55	792	22.0

From this table it is clear that the *time* as well as the *number* of repetitions necessary for learning increases much faster than the length of the list to be learned, and that neither of these measures of memory ability increases in simple arithmetic or geometric progression. There is, relatively speaking, an enormous jump in the number of readings needed as we go from seven to ten syllables; but after this the increase is roughly constant and at the rate of about 50 per cent from each syllable list to the next following. If we disregard the again comparatively large increase in time of learning from the seven-word to the ten-word list, the time increment also is roughly constant and at the rate of about 100 per cent. It is interesting to compare the effort involved in learning twelve, twenty-four, and thirty-six syllables, as these lists are in the ratio 1:2:3. Taking the "average time per syllable" for a list of twelve, i.e., 6.8 seconds per syl-

lable, as our standard, the list of twenty-four syllables requires two and one-half times as much time per syllable as the list of twelve, and the list of thirty-six syllables requires three and two tenths as much time. The absolute increase is about 11 seconds per syllable, on the average, from twelve to twenty-four syllables and about 5 seconds per syllable from twenty-four to thirty-six; hence it is clearly harder to go from twelve to twenty-four syllables than from twenty-four to thirty-six.

As the memory task increases in length, not only do we get an increase in the number of readings and in the time required—which is to be expected—but a marked increase in the *time per syllable* as well. This suggests that the greater the number of associations required, the greater is the effort which must be expended on each association; and that learning a long list is not, therefore, simply a matter of adding on *more* syllables to a short list. The moral for the student lies in recognizing that he must expend not only more time and effort in learning a long than in learning a short lesson, but more time and effort per item or unit as well.

A partial explanation of this finding would seem to lie in the need for linking together parts or sections of the longer lists as well as of tying together the separate syllables or elements. It is highly probable, too, that in the longer lists the later-formed associations serve to confuse and to be confused by those which come earlier; and as a result *all* require better fixation, i.e., more learning. This interference of one set of associations with another set is called *retroactive inhibition* when it extends in the backward direction, and *proactive inhibition* when it extends in the forward direction. Both kinds of confusion have been studied by psychologists, especially the retroactive effects which result when a memory task is followed immediately by other memory tasks. Experimental data show that the confusion effect on recall is greatest when the activity which is interpolated between fixation and recall involves the learning of material closely similar to the original lesson. One is likely, for example, to remember more of a French vocabulary if he studies algebra in the interval between fixation and recall than if he studies Latin in the interpolated interval. Experiments have shown, too, that forgetting, although it occurs during sleep, is less rapid than when we are awake. Interference and association confusions seem to be better explanations of memory loss than decay resulting from the passage of time. Of course, many other factors enter into and condition forgetting. There is a large body of experimental literature on the subject.

By way of contrast with the data from nonsense material, it is interesting to compare some results got by Ebbinghaus with meaningful subject-matter. In seven tests, each test comprising six stanzas from

Byron's *Don Juan*, Ebbinghaus found that he required on the average about eight repetitions to reach his standard of one errorless recitation. Each stanza contained about eighty syllables. On the basis of the results given in Table VIII, Ebbinghaus estimated that he would probably have required about seventy to eighty repetitions to learn eighty nonsense syllables up to his standard of one errorless recitation. A comparison of these figures for meaningful and meaningless material gives us a ratio of 10:1 in favor of the meaningful, and furnishes, in Ebbinghaus's words, "an approximate numerical expression for the extraordinary advantage which the combined ties of meaning, rhythm, rhyme, and a common language give to material to be memorized."

Objection has been raised to Ebbinghaus's standard of learning, namely, the time and the number of repetitions necessary to reach the first errorless reproduction. Impetuous subjects, it is said, will attempt to reproduce a list before it has been well learned, and hence run the risk of becoming confused and discouraged; while cautious individuals will study longer than is necessary, and hence tend to "overlearn." This "error of variable standard" can be controlled in a practical way by requiring a subject to attempt reproduction early—but not too early—in his learning, thus making sure that the first errorless recitation is really the first reproduction of which he is capable. With trained subjects little difficulty arises on this account. Probably no one since Ebbinghaus has been better trained or more careful than he; and accordingly his results are highly reliable, although based upon data from only one person.

4. THE AMOUNT OR DEGREE OF LEARNING AND RETENTION

The reader will readily understand that the results quoted in the last section are limited in scope, since in every case Ebbinghaus only learned his lists up to the point of one correct reproduction. Ordinarily we memorize to a considerably greater degree than this, and in general the slower the learner, the more often and the more painstakingly does he go over his task. What is the influence of this "overlearning"² upon retention? Ebbinghaus undertook to answer this question in the following experiment: He read lists of sixteen syllables in exactly the same way and at the same rate, except that the number of readings varied from eight to sixty-four. This meant, of course, that some of the lists were overlearned to a very high degree. Twenty-four hours later he studied these same lists until he could just recite them once. He then figured the percentage saved in re-learning each list—i.e., he found how many repetitions *less* were required

² According to Ebbinghaus's standards, any learning over and above that necessary to obtain one correct reproduction constitutes overlearning.

to relearn than to learn twenty-four hours before, and what per cent this saving was of the original learning time. This method of studying memory was devised by Ebbinghaus and is known as the "saving or savings method." Following are some of his results based on several lists:

TABLE IX

Number of readings on the first day	8	16	24	32	42	53	64
Per cent saved in relearning the lists 24 hours later	8	15	23	32	45	54	64

Evidently each original repetition brought about a saving of almost exactly 1 per cent in the relearning necessary on the following day. The remarkable uniformity of these results suggest that 100 readings on the first day should logically make it possible to reproduce a list on the next day without any additional learning—the saving would be 100 per cent! But Ebbinghaus found that it was impossible for him to read through a list 100 times without serious lapses of attention, fatigue, and drowsiness. This indicates that the regular saving of 1 per cent for each previous day's reading will eventually break down as the learner's "physiological limit" is reached; and that beyond this point further repetition will be of little or no value.

The clear implication of this experiment is that degree of retention depends upon the amount of work done (overlearning), as well as upon the amount to be learned. In memorizing nonsense syllables the strength of the connections formed, as shown by the saving in relearning, is, up to a certain point, roughly proportional to the amount of study. This fact throws light upon the widely held view that the slow learner retains better than the fast learner. Perhaps the truth of the matter is that the slow learner does more overlearning than the fast learner, and hence saves more in relearning. In this connection the following data from Radosavljevich³ are of interest. This experimenter compared children with adults in the learning of twelve-syllable nonsense lists and the relearning of the same lists after twenty-four hours. His results were as follows:

TABLE X

	No. Readings in First Learning	No. Readings Needed to Relearn 24 Hours Later	Per Cent of Saving
Adults	20	6	70
Children	42	7	83

³ A later investigator (1907), who repeated much of Ebbinghaus's work using more subjects. See Ladd and Woodworth (ref. 11).

The children took twice as many readings to learn the lists originally, but they saved a considerably larger per cent in relearning than did the adults. Very probably this is what happens in the case of the slow but sure learner, and accounts for his apparent superiority in retention over the rapid learner. When the *amount* or *degree* of learning is kept constant the quick learner retains better than the slow learner.

Of the many studies of degree of learning and retention which have followed those of Ebbinghaus, one of the most thorough is that of Luh. This experimenter had ten subjects memorize lists of twelve nonsense syllables to four degrees or stages of completeness: 100 per cent, 150 per cent, 67 per cent, and 33 per cent. In 100 per cent learning each subject memorized his lists up to the point of *one* errorless recitation; in 150 per cent learning he was allowed as many repetitions as in 100 per cent learning plus a bonus of half as many again; in 67 per cent learning he was given two thirds as many repetitions as in 100 per cent learning; and in 33 per cent learning, one third as many. Retention was tested by three methods: *written reproduction*, *recognition*, and *reconstruction* (for description, see later p. 113). Luh's results showing the amounts retained after three time-intervals are given in Table XI.

TABLE XI

Memory Method	Amount Retained after Three Time-Intervals		
	4 hours	1 day	2 days
Written reproduction			
150% learning	82%	39%	31%
100%	65	46	40
67%	66	42	25
33%	43	26	14
Recognition			
150% learning	93%	83%	73%
100%	92	78	79
67%	85	74	62
33%	55	46	26
Reconstruction			
150% learning	91%	43%	44%
100%	75	49	44
67%	65	57	32
33%	48	26	20

On the whole these data support Ebbinghaus insofar as they indicate that increasing the repetitions (up to a certain point) leads to better retention. The "diminishing returns" from overlearning, commented upon

by Ebbinghaus, are clearly shown in Luh's records, since after intervals of one day and two days better results are obtained from 100 per cent than from 150 per cent learning. However, the loss from overlearning appears much sooner in Luh's than in Ebbinghaus's results. From Table VIII we find that it took Ebbinghaus thirty repetitions to learn sixteen nonsense syllables up to the standard of one correct recitation. If we take thirty-two repetitions by Ebbinghaus (see Table IX) as roughly equivalent to Luh's 100 per cent learning, it would seem that Ebbinghaus overlearned his material up to 100 per cent (sixty-four repetitions) without any diminishing returns appearing. This same result—a failure to get diminishing returns from overlearning—also appears in the work of Krueger, who checked some of Luh's results. Krueger had twenty subjects learn lists of twelve monosyllabic words (nouns) presented at 2-sec. intervals on a rotating drum apparatus. Degree of learning was varied from 100 per cent learning (no overlearning) up to 200 per cent learning (100 per cent overlearning); and retention was tested after one, two, four, seven, fourteen, and twenty-eight days. Krueger found that overlearning up to 50 per cent (150 per cent learning) led to an increase in retention almost directly proportional to the additional effort. But no evidence was found of large diminishing returns with overlearning even for the longer periods between learning and recall. Krueger's results, therefore, substantiate Ebbinghaus's and are opposed to Luh's insofar as they indicate little or no harmful effect from excess learning.

It should be pointed out that Luh's findings are not strictly comparable to those of Ebbinghaus or of Krueger. Luh used different methods of presenting his learning material and of measuring memory retention; and his lists were shorter than those of Ebbinghaus—12 *versus* 16 nonsense syllables. Again, Luh's results are based upon the average of 10 persons while Ebbinghaus's findings are for himself alone. Krueger's lists were of the same length as those of Luh (*viz.* 12) but were meaningful words instead of nonsense syllables. Hence a part at least of the apparent discrepancy in their results may be attributed to differences in method and differences in materials.

On the whole, except in the case of nonsense syllables, there is little experimental evidence of memory loss attributable to overlearning. Probably there is a point beyond which the time spent by a student on German grammar or geometrical propositions represents wasted effort. But this lack of return is perhaps due to boredom (*i.e.*, to indifference or loss of interest) and fatigue rather than to the fact that the added study is in itself of no value. Summarizing our positive findings from these experiments, it seems evident that overlearning up to a certain point (to be located experimentally) definitely increases retention.

5. THE TIME-INTERVAL BETWEEN LEARNING AND RECALL

No doubt many of us have had the experience of discovering that a poem memorized long ago, although it cannot be recited now, can be relearned in a much shorter time than a new poem of the same length and the same approximate difficulty. Fortunately, this quicker relearning holds for most of the facts acquired in school for which retention appears to be almost zero. Indeed, the ability to "get back" material once studied but now forgotten seems oftentimes to be the only real justification for having learned it at all. It is common experience, too, that the more recently we have studied a topic the fresher it tends to be. Just how is this forgetting related to the passing of time? Do memories dim gradually and progressively, or is a large amount lost comparatively quickly, some details with possibly a vague outline remaining for a longer time? Ebbinghaus attacked this problem in what is probably his most important, and is certainly his best-known, experiment. Eight lists of thirteen nonsense syllables each were learned up to the point of *two* errorless recitations. Then, after a lapse of twenty minutes, the lists were again taken up and studied as before until each list could be recited *twice* without error. The actual time saved in relearning was expressed as a per cent of the original learning time (the *saving method*). To illustrate the procedure, it took Ebbinghaus on the average about eighteen minutes (1,080 seconds) to learn twelve lists of thirteen nonsense syllables each up to his new standard of two correct reproductions. After a lapse of twenty minutes it required only about eight minutes (498 seconds) for him to relearn the same lists. This means there was a saving of about $10/18$, or 56 per cent,⁴ of the time which Ebbinghaus required to learn the lists originally. Fifty-six per cent measures the retention or the amount which "stuck" over a period of twenty minutes and shows that the forgetting during this time was 100 minus 56, or 44 per cent. Other lists of thirteen syllables were learned by Ebbinghaus in the same manner and relearned again after periods of one hour, nine hours, one day, two days, six days, and thirty-one days. The per cent of saving in each case was calculated after the intervals stated above. In memorizing his lists and then relearning them, Ebbinghaus was careful to distribute his learning so as to prevent any one period from being unduly favored by practice. Table XII summarizes his principal results.

If we draw a graph of these results, laying off the time-intervals between learning and relearning along the baseline and the per cents re-

⁴ This percentage was determined more accurately to be 58 per cent by Ebbinghaus when all of his data were considered.

TABLE XII

<i>Interval between Learning and Relearning</i>	<i>Percentage of Work Saved (Retention)</i>	<i>Loss Due to Forgetting</i>
20 minutes	58	42
1 hour	44	56
9 hours	36	64
24 hours	34	66
2 days	28	72
6 days	25	75
31 days	21	79

tained on the vertical axis, we get Ebbinghaus's *Curve of Retention* (Fig. 23). This graph, which has been a classic in psychology for all of fifty years, has the general character of an inverse logarithmic relationship.⁵ It indicates that, after the large initial drop from twenty minutes to two days, forgetting proceeds more and more slowly until there is very little difference between the loss after ten days and after thirty days. Presumably a certain minimum sticks almost indefinitely; and it is just this minimum which tips the balance against the uneducated man in favor of the educated man long out of school.

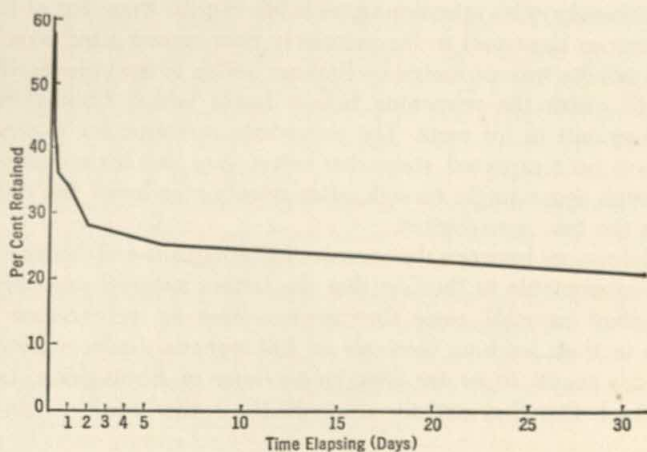


FIG. 23
CURVE OF RETENTION (EBBINGHAUS) FOR NONSENSE SYLLABLES AFTER
VARIOUS TIME INTERVALS

⁵ Its equation is $b = \frac{100k}{(\log t)^c + k}$ in which b = per cent retained, t = time-interval elapsing, and k and c are constants.

Later investigators working with many subjects and with meaningful material as well as with nonsense syllables have obtained results different in some respects from those of Ebbinghaus, but not markedly different. The following table from Radosavljevich gives the records from several subjects both for nonsense syllables and poetry.

TABLE XIII

<i>Interval between Learning and Relearning</i>	<i>Percentage of Retention for Nonsense Syllables</i>	<i>Percentage of Retention for Poetry</i>
5 minutes	98	100
20 minutes	89	96
1 hour	71	78
8 hours	47	58
24 hours	69	78
2 days	61	67
6 days	49	42
14 days	41	30
30 days	20	24
120 days	3	

In Table XIII, forgetting proceeds at a slower rate than was found by Ebbinghaus, but the general trend in the two experiments is much the same. Radosavljevich's retention curve is less regular than that of Ebbinghaus owing in large part to the extremely poor record after eight hours. This irregularity was attributed by Radosavljevich to the unfavorable time of day at which the relearning fell—a factor which Ebbinghaus had guarded against in his work. The immediate retention for poetry is, as might have been expected, somewhat better than that for nonsense syllables, though, surprisingly enough, after twenty-four hours the difference between the two is negligible.

The differences between the records of Ebbinghaus and Radosavljevich are also attributable to the fact that the latter's subjects probably overlearned their material, since they were neither so well trained nor so uniform in their learning methods as Ebbinghaus. Radosavljevich considered his results to be far superior to those of Ebbinghaus, but it is doubtful whether they actually discredit the earlier work.

6. INFLUENCE OF THE METHOD OF MEMORIZING UPON RETENTION

Besides the matter of individual differences and overlearning, the retention of nonsense syllables depends to a considerable degree upon the *method* employed in measuring retention. This has been clearly shown by Luh in an experiment in which five relearning methods were compared:

anticipation, *relearning* (or *saving*), *written reproduction*, *reconstruction*, and *recognition*. Throughout Luh's work the presentation of nonsense syllables was by means of a rotating drum apparatus which exposed the twelve syllables one at a time through an aperture in a screen. In the *anticipation* method, the subject attempted in his reproduction to anticipate each syllable before it was shown; the *relearning* method is the saving method of Ebbinghaus previously described; in *written reproduction* the subject simply wrote down all of the nonsense syllables which he could remember; in *reconstruction* he was given the twelve syllables, each on a separate slip of paper, and was instructed to rearrange them in the original order of presentation; and in *recognition* he attempted to select from twenty-four syllables the twelve which he had seen. Luh's results are shown in Table XIV and his retention curves in Figure 24.

TABLE XIV

Method	Percentages Retained after				
	20 mins.	1 hour	4 hours	1 day	2 days
Anticipation	68	50	39	18	10
Relearning (saving)	75	66	55	52	48
Written reproduction	88	82	61	39	27
Reconstruction	92	90	75	51	39
Recognition	98	95	93	75	72

The striking fact in Table XIV is the discrepancy in percentages retained after the five given intervals. Clearly the method employed in recording amounts retained is sufficient in itself to make the losses after various time-intervals decidedly different in size. The hardest method is anticipation, which is almost pure recall, with a minimum of "cues"; the easiest method is recognition, in which the loss after two days is slightly less than the loss after twenty minutes by the anticipation method. This result, namely, that recognition is decidedly easier than recall, has been found by other investigators.

The percentages retained as found by the relearning method are shown in Table XV. These are the only data directly comparable to those of Ebbinghaus (see Table XII), since Ebbinghaus employed only the relearning (i.e., saving) method. A comparison of the percentages saved for four time-intervals as found by Ebbinghaus and Luh is as follows:

TABLE XV

	Percentages Saved after			
	20 mins.	1 hour	1 day	2 days
Ebbinghaus	58	44	34	28
Luh	75	66	52	48
Difference	17	22	18	20

Ebbinghaus's results run regularly from 18 to 20 per cent below those of Luh for comparable time-intervals. Part of this difference is due to

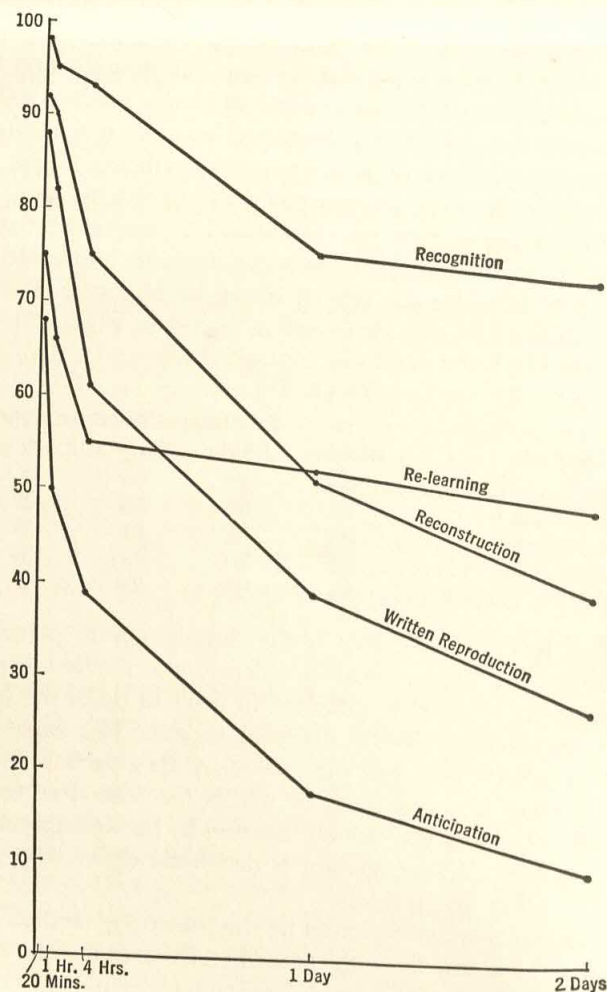


FIG. 24

CURVES OF RETENTION FOR TWELVE NONSENSE SYLLABLES LEARNED BY FIVE DIFFERENT METHODS

Percentages retained are laid off on the vertical axis; times between learning and relearning on the horizontal axis. (From Luh, 1922.)

the fact that Ebbinghaus learned 13 nonsense syllables, while Luh's subjects learned only 12. From Table VIII we know that the addition of two extra syllables (the jump from 10 to 12) increases the number of readings necessary for learning from 13 to 17; hence an increase of one syllable might be expected to have some effect, at least, on the saving from re-

learning. But the greater part of the difference is undoubtedly due to the method of presentation. Luh's presentation was rigidly controlled, each syllable being shown as an isolated unit so that connections must have been almost 100 per cent in the forward direction. On the other hand, Ebbinghaus's lists, which were spread out before him, offered ample opportunity for backward as well as forward associations and for associations with the list as a whole (see later, p. 118). If learning was easier in Ebbinghaus's procedure than in Luh's, the enforced overlearning of the latter's subjects as compared with Ebbinghaus would inevitably have made forgetting less rapid.

7. THE RETENTION OF SCHOOL SUBJECTS

Leaving now the more carefully controlled laboratory experiments on memory retention, let us consider some investigations which have sought to discover the permanence of subject-matter learned in school. Table XVI shows the per cent retention of certain school subjects at the end of the course and after one and two years.

TABLE XVI
RETENTION OF VARIOUS SCHOOL SUBJECTS AT END OF COURSE AND
AFTER ONE OR TWO YEARS

(From Pressey, p. 393)

	<i>Per Cent of Perfect Score on Test Used</i>		
	<i>End of Course</i>	<i>After 1 Year</i>	<i>After 2 Years</i>
Botany (college) ¹	.68	.21	.16
Zoölogy (college) ¹	.78	.39	.30
Psychology (college) ¹	.70	.24	.19
Algebra (high school) ²	.87	.56	...
Latin (high school) ¹	.82	.60	...
Chemistry (high school) ¹	.63	.47	.33
History (eighth grade) ³	.71	.56	...
Average	.74	.43	.24

¹ Results based upon objective tests.

² Results based upon Regents' examination.

³ Results based upon essay-type questions.

These percentages undoubtedly represent a kind of optimum, as many of the poorer and less interested students presumably dropped out and were not retested, especially after one and two years. It is doubtful whether the per cents recalled of high school algebra and Latin would be as large as shown here for students in general. Differences in percentages retained as shown in Table XVI are probably due to variations in initial ability, difficulty of course, effectiveness of instruction, and the like rather than to

differences in subject matter. On the whole, the students tested showed a retention of about 75 per cent at the end of the course and of about 25 per cent after two years.

Results very much like these were found in a more intensive study of the retention of lecture material by college students after various time intervals. In several groups of subjects, it was found (ref. 9) that at the close of a forty-minute lecture students can reproduce on the average about 62 per cent of the material just presented; after three or four days, about 45 per cent is remembered; after one week, 35 per cent; after two weeks, 31 per cent; and after eight weeks, 24 per cent. The retention curve plotted from these data falls off in much the same fashion as that of Ebbinghaus for nonsense syllables: i.e., the loss is rapid at first and progressively slower as time goes on.

Results of much practical importance have been obtained in studies of "school forgetting." Over a summer vacation, elementary school children show an increase in ability to read and a loss in spelling, arithmetic, and history. Here the obvious explanation is that children exercise (over-learn?) reading skills in their everyday activities during the summer, but ignore and hence forget the more "scholastic" subjects. This tendency to forget what is not used holds, of course, for adults as well as for children. The physician forgets his poetry, the engineer his history, and the business man nearly everything except his economics. The relative "permanence" of material learned by different methods of instruction has been studied exhaustively by educational psychologists.

Since forgetting is much faster during the period immediately following learning, obviously the thing to do is to review early and often. In this way what one has learned will be held, so to speak, above the "memory threshold." It must be kept in mind that Ebbinghaus barely learned his material in the first place. We have found already how enormously retention is affected by overlearning. Therefore, when a memory lesson is exceedingly well learned to begin with, forgetting must proceed at a very much slower rate than that exhibited by Ebbinghaus's curve (see Fig. 23). The effect of repeated learning will be considered in the next section.

8. THE EFFECT OF REPETITION AND FREQUENT REVIEW UPON ABILITY TO RETAIN

We have seen in the last sections that material just barely learned tends to be forgotten very rapidly at first and then more slowly as time goes on. Also, it appears clearly in Ebbinghaus's and Luh's results (Tables IX and XI), that the amount forgotten depends upon the method of recording

retention, as well as upon the labor expended in the original learning. What effect, we may now ask, does repeated learning (review) have upon retention? Ebbinghaus attempted to answer this question in the following experiment. First, series of twelve, twenty-four, and thirty-six nonsense syllables as well as stanzas from Byron's *Don Juan* were learned. These were then relearned at the same hour on six successive days, each time up to the standard of one correct recitation. Table XVII shows the number of repetitions required at each period and the per cent saved over the first day's record.

TABLE XVII

<i>No. of Syllables in Series</i>	<i>No. of Repetitions Which (on the Average) Were Necessary for Relearning the Series on Successive Days; also Per Cent Saved on Suc- cessive Days over the First Day's Record</i>					
	Days					
	I	II	III	IV	V	VI
12 Number of repetitions	16.5	11	7.5	5	3	2.5
% saved over first day	...	33	55	70	82	85
24 Number of repetitions	44	22.5	12.5	7.5	4.5	3.5
% saved over first day	...	49	72	83	90	92
36 Number of repetitions	55	23	11	7.5	4.5	3.5
% saved over first day	...	58	80	86	92	94
1 stanza <i>Don Juan</i>						
Number of repetitions	7.75	3.75	1.75	.5	0	0
% saved over first day	...	52	77	94	100	100

The important fact in Table XVII is that the necessary amount of relearning becomes progressively less and less on each succeeding day. In other words, the more exercised or older associations are forgotten more slowly than the less exercised or newer associations. On each succeeding day, Ebbinghaus brought his learning up to the standard of the previous day—one correct recitation. If this means that the associations found were left in the *same* condition at the end of each day's learning, then logically we should expect the loss to be about the same from one day to the next, no matter how often the learning was repeated. To use an analogy, forgetting according to this view might be conceived of as the running down of a clock which must be rewound to the same point each day in order to bring it back to its original efficiency. So if it took eleven repetitions on the second day to bring the twelve nonsense syllables up to their original strength on the first day, it should take eleven repetitions to bring them up to the same standard on the third day, the fourth day, the fifth, and the sixth. On the contrary, we find that the loss due to for-

getting became less and less after each learning period. This can mean but one thing, namely, that connections are more strongly established the more often they are exercised. So the multiplication table, the months of the year, and the names and relations of common coins are rarely forgotten except by very old persons or by those suffering from mental disease.

It is interesting to note in Table XVII that the repetitions, when the learning was repeated at stated intervals, were relatively fewer for the longer lists. This is owing to the fact that many more repetitions were needed originally to fixate the longer series; hence from the beginning the associations in these lists were the more firmly established. The saving, for example, in relearning thirty-six nonsense syllables after twenty-four hours was 58 per cent as against a saving of 33 per cent for twelve nonsense syllables after the same interval. But nearly three and one-half times as many repetitions were needed to learn the longer list in the first place.

An important discovery made by Ebbinghaus in the present connection has to do with the relative value of *distributed* versus *concentrated* learning. Ebbinghaus found that sixty-eight successive readings of a twelve-syllable list made it possible for him to relearn the same list on the following day after only seven repetitions. When his readings were spaced over three days, the same result was not attained until the fourth day—but this time only thirty-eight readings were necessary! In other words, a judicious distribution of effort over three days produced a result which when learning was packed into a single day required *twice* as much labor. The superiority of spaced over massed learning has since been verified by other experiments; more will be said about it in Section 10.

9. THE CONNECTIONS FORMED IN LEARNING

Do the connections formed in learning run forward serially from one term to the next following only, or do they skip terms in the forward direction, and even sometimes in the backward direction? What is the relative strength of these different associations, if they can be shown to exist?

In serial learning, it often seems that in addition to simple 1, 2, 3, or *a, b, c* forward associations, other connections are set up in various ways within the material to be learned. In the experiments of Ebbinghaus hitherto described, repetition always proceeded regularly from one syllable to the next following. Hence, it would appear that only serial connections could be made, and that these connections must be formed in the

forward direction only. Are there any other associations formed in rote learning?

Ebbinghaus set out to answer this question first by learning lists of sixteen syllables up to the point of one correct recitation, and then twenty-four hours later by learning "derived" lists made up in various ways from the original lists. If we designate the syllables of an original list of numbers, 1, 2, 3 . . . 16, a sample derived list might read:

1, 3, 5, 7, 9, 11, 13, 15, 2, 4, 6, 8, 10, 12, 14, 16.

Such a derived list would clearly benefit from forward associations between one term and the next following if such associations had been formed in learning the original list. Also, if this is true, it should be easier to relearn such a derived list than an entirely new one. Another derived list might skip two syllables in the forward direction:

1, 4, 7, 10, 13, 16, 2, 5, 8, 11, 14, 3, 6, 9, 12, 15,

thus giving more remote forward associations a chance to operate if present. Still other derived lists drawn up by Ebbinghaus skipped three, four, and up to seven intervening syllables. In addition to lists like these, others were constructed in which the syllables were reversed in regular order from 16 to 1, or were simply jumbled up or arranged in chance order. Thus forward associations, both near and remote, backward associations, and any other incidental connections were given a chance to operate in the relearning if present in fact.

The average saving in terms of repetitions in learning these derived lists twenty-four hours after the original lists had been learned is shown in Table XVIII.

TABLE XVIII

Saving in relearning original lists unchanged	33%
Saving in learning lists derived by skipping 1 syllable	11%
2 syllables	7%
3	6%
7	3%
Saving in learning reversed lists	12%
reversed lists, 1 syllable skipped	5%
lists, syllables arranged by chance	.5%

The percentages of saving shown in the table are rather small in several cases, but they are based upon six or more lists and are statistically reliable. It seems clear that in learning these lists connections were formed not simply from one syllable to the next following, but to the second, the third, and even more remote terms. To quote Ebbinghaus: "As a result of the repetition of the syllable series, certain connections are established

between each member and all those that follow it. These connections are revealed by the fact that the syllable pairs so bound together are recalled to mind more easily and with the overcoming of less friction than similar pairs which have not been previously united." The more remote the terms, the weaker the connection tends to be. Only in the case of derived lists in which the syllables are simply thrown together by chance or reversed with one syllable skipped is there no real saving.

An explanation of "backward," "remote forward," and "incidental" association is to be found, perhaps, in the learning method employed by Ebbinghaus rather than in any reversible neural mechanism. With his lists of nonsense syllables spread out before him, it was almost impossible for Ebbinghaus to refrain from glancing backward and forward as he read slowly through a list. Hence his method might account for the adventitious connections which showed up later on in learning the derived lists. It seems likely that much of the benefit obtained from overlearning is brought about by a strengthening of the essential forward associations together with a gradual weakening of indirect and useless connections. With continued repetition, bonds which might easily cause confusion and interference when learning is imperfect are minimized in favor of the more necessary direct ones.

10. ECONOMY AND EFFICIENCY IN MEMORIZING

In Section 2 it was pointed out that Ebbinghaus and many later psychologists have been especially interested in the fixation process because without adequate registration there can be little effective recall and recognition. One can probably do little to improve his physiological retentivity other than to build up his physical and emotional well-being; apparently some minds are natively more impressionable than others, just as some sense organs are natively more sensitive than others. Various schemes and methods have been utilized, however, for improving memory-fixation, and for making recall and recognition more efficient. Some of these will be outlined in the present section.

1. *Spaced versus Massed Learning.* When Ebbinghaus massed his learning of nonsense syllables into a single day, he obtained a result (p. 118) which required only one half as much effort when distributed over three days. The superiority of distributed or spaced learning over concentrated or massed learning appears in laboratory studies as well as in studies of learning in school. Both with respect to amount retained and accuracy of recall, spaced learning of three-place numbers has proved to be better than massed learning. In an experiment dealing with memory for a prose passage of 1,500 words, it was found that the material was better re-

membered after four readings spaced one day apart than after four readings in immediate succession. Spaced learning is but little—if any—better than massed learning when the memory lesson is easy—there is little value in repeating an easy task. But spaced learning has a decided advantage over massed when the material is difficult, as it is here that repeated attack is most necessary and most effective. The superiority of distributed learning is due, perhaps, to many factors: (a) when the learning periods are comparatively short, there is less tendency to fall into a rut and repeat errors; (b) frequent review maintains the freshness and recency of the memory task; (c) there is better attention and less boredom when the learning period is not too extended.

The finding that spaced learning is better than massed is important for school practice since it indicates that frequent repetition and review are more effective and more saving of time and labor than learning concentrated into one grand effort. It is sometimes true, of course, that a clever student will be able to “cram” enough of what has been taught during the term to pass the next day’s examination. But he will almost surely retain less of the content of the course than his more sagacious brother who has distributed his learning throughout the term. And so in the matter of retention, at least, virtue is rewarded!

2. *Whole versus Part Learning.* When material of any length is to be memorized, it is generally a better plan to study it straight through from beginning to end (“as a whole”) than to split it up into bits and learn each part separately. Nonsense syllables as well as poetry and prose are learned more efficiently by the “whole” method; and retention over intervals of time is better. Poetry learned by the whole method, for example, was recalled 160 per cent better at the end of two years than poetry learned by the part method.

The reason why the whole method is better than the part is that associations in whole learning are formed in the order and direction in which the learner wants them to function eventually. It is not necessary as in part learning to “tie together” sections often artificially separated. The whole method would seem, too, to be the more intelligent plan of attack. It is probable that the better students work out some plan of “whole” learning (finding relations and implications) for themselves, while the poorer students waste time on details and unrelated bits of knowledge.

When a lesson is long and difficult, a combination of the whole and part methods is the most efficient mode of attack. The material is first surveyed as an entirety in order to give perspective and provide logical relations in terms of which the ideas presented may be integrated. Intensive study is then directed selectively to those parts of the memory-lesson which are most difficult. Breaking up a long lesson into “natural”

divisions prevents the learner from being overwhelmed, and hence emotionally disturbed, by the size of his job.

3. *Active versus Passive Learning.* Closely related to the question of the best distribution of learning effort and plan of attack is that of the value of recitation during the learning itself. If a memory task is studied by first reading it through and then reciting it to oneself, rereading those parts not clearly remembered, again reciting and so on, retention will be better than if one reads and rereads without recitation. Table XIX shows the results of an experiment in which both nonsense syllables and meaningful material were employed. The subjects were eighth-grade pupils. Evidently, recitation is a decided aid to recall and the more recitation the better the recall.

TABLE XIX
(From Gates, 1917)

	<i>Nonsense Syllables, Per Cent Remembered:</i>		<i>Five Short Biographies Totaling about 170 Words, Per Cent Remembered:</i>	
	<i>Immediately</i>	<i>After 4 Hours</i>	<i>Immediately</i>	<i>After 4 Hours</i>
All time given to reading	35	15	35	16
$\frac{1}{2}$ time given to recitation	50	26	37	19
$\frac{2}{3}$ time given to recitation	54	28	41	25
$\frac{3}{4}$ time given to recitation	57	37	42	26
$\frac{4}{5}$ time given to recitation	74	48	42	26

Students can greatly aid their retention by training themselves in the simple technique of reading and reciting, often prompting themselves by reference back to the lesson when necessary. The value of self-recitation for recall and recognition lies in the fact that in reciting to himself the learner is reacting actively to the task. This active search for connections and implications is often called the "will to learn," and is most effective in memory work. In reading or listening to a lecture the student is forced into a passive rôle. In self-recitation, instead of sitting quietly hoping that somehow the lesson will record itself photographically upon his "mind," the student aids the recording process (fixation) by taking an active part in it. Writing abstracts of what one has read and discussion of this material with others aids recall by providing richer connections and associations with the material.

4. *Overlearning and Review.* We have seen in Section 4 how "overlearning" (i.e., learning beyond the bare reproduction level) strengthens retention. Frequent and strategically placed reviews also aid recall and recognition. Reference to Figure 23 will show that the loss through for-

getting is most rapid at first, becoming slower and slower as time goes on. Retentivity is helped, then, by reviewing early and often; early to catch the first abrupt sag in the retention curve, and often to prevent later and recurring sags. There is much evidence to support this contention. When a forty-minute lecture was followed immediately by a five-minute review, recall after two months was 50 per cent better than when no immediate review was given.

5. *Memory "Training."* Since failure to remember is such a universal complaint, it is small wonder that various systems claiming to "train" or improve one's memory have been devised. All such systems depend essentially upon the building up of associations between the facts to be learned and the facts already known. When the associations formed are meaningful and relevant (when new facts are hung, so to say, upon the peg of some known fact) the system is helpful. But when extraneous and meaningless associations are advised, it often becomes more difficult to remember the system than to learn the lesson in the first place.

There are several techniques for aiding memory which are useful and hence are worth mentioning.

(a) In memorizing number series, telephone or street numbers, say, *grouping* is helpful. Stanzas, rules, dates, paragraphs, sections, and the like which represent logical groupings should be utilized whenever possible. Often the material will fall into "natural" groups—the succession 25, 35, 45, for instance, is easier than 27, 36, 48. When *rhythm* is added to the grouping, as in music or poetry, recall is still further aided.

(b) Attention to *meanings* and the formation of *secondary associations* helps recall and recognition. If one has good visual imagery, one may aid recall of three specific facts by means of a diagram—by thinking of the three as perched upon the apices of a triangle. Or he may rule off a square into smaller squares (in his "mind's eye") and think of each section as a pigeonhole into which an article is placed. Objects and facts which can be related to important happenings in one's life are more likely to be recalled. Thus, a man may remember the age of a neighbor's child whose age is the same as his own child but forget other children's ages; a boy remembers football signals but forgets his algebra assignment; a girl remembers her "date" but forgets to wash the dishes and so on. Meanings are important aids to memory. Dates in history are better remembered as milestones marking off social, political, or economic periods than as times at which a certain battle took place. Interest and utility, by investing many otherwise tedious statistics with meaning, enhance their memory-value.

(c) "Crutches" or memory aids are often useful when the facts have no inherent organization in themselves; though many such aids are proba-

bly more of a hindrance than a help. A far-fetched association or idea brought in to aid recall may itself be forgotten; and an ambiguous aid though it may itself be recalled may not revive what it is supposed to revive and may even hinder such revival. The well-known doggerel "Thirty days hath September, etc." is helpful only if one can remember the doggerel! The same is true of the verse "In 1492 Columbus sailed the ocean blue," "oysters *r* in season" and many others. Counting on one's fingers may be definitely useful at first, but it becomes a hindrance and a nuisance if persisted in. When crutches work it is because of their novelty and humorous appeal rather than because of their inherent value.

Probably more people, especially older people, complain of their failure to remember names than of any other memory loss. Part of the trouble here may be due to loss of plasticity with age, but the failure to form adequate secondary associations is probably more important. As one gets older the increase in number of acquaintances puts an inevitable burden upon one's memory and the number of friends with the same or nearly the same name makes for confusion.

(d) Failure to remember a name is often due to poor fixation in the first instance—a name may have been mumbled or not clearly heard upon introduction. A hint on how to improve name fixation is provided by those in whose work it is important to remember names and faces. Such persons upon introduction to a stranger repeat the name with some emphasis and find occasion to use it later on. Secondary associations of names with the characteristics of persons will also help, and are carefully noted by many people. One may remember that Mr. Fox has a long, thin face and a prominent nose; that Mr. Johnson is tall and blond; that Miss Redding has red hair; that Mr. Lee is a Southerner and so on. For those whose memory for names is notoriously poor, writing down the names of people one often forgets with brief descriptions of each, age, appearance, occupation, family connections, and outstanding traits, will help. More and better associations is still the best advice that the psychologist can give the person with poor memory.

11. SUMMARY OF EBBINGHAUS'S WORK

Leaving our digression into the practical application of experimental studies in memory to affairs of everyday life, let us attempt to summarize in a few words Ebbinghaus's chief contributions to experimental psychology. Certainly we must list (1) his introduction and use of quantitative methods in the study of learning and forgetting; (2) his measurement of the factors governing fixation, retention, and recall in verbal learning; and (3) his invention of nonsense syllables. Ebbinghaus's memory methods

are today standard procedures in the psychological laboratory. His main results may be accepted substantially as he left them. Of his invention of nonsense syllables Titchener remarks (ref. 21, p. 380) "It is not too much to say that the recourse to nonsense syllables, as a means to the study of association, marks the most considerable advance, in this chapter of psychology, since the time of Aristotle." In short, Hermann Ebbinghaus was the founder of the experimental (quantitative) study of memory.

BIBLIOGRAPHY

1. Britt, S. H. Theories of retroactive inhibition. *Psychol. Rev.*, 1936, 43, 207-216.
2. Ebbinghaus, H. *Memory, a contribution to experimental psychology*, 1885. Translated by H. A. Ruger and C. E. Bussenius. New York: Teachers College, Columbia University, 1913.
3. English, H. B., Welborn, E. L., & Killian, C. D. Studies in substance memorization, *J. gen. Psychol.*, 1934, 11, 233-260.
4. Gillette, A. L. Learning and retention. *Arch. Psychol.*, New York, 1936, 28, No. 198.
5. Hovland, C. I. Experimental studies of rote learning theory. IV. Comparison of retention following learning to same criterion by massed and distributed practice. *J. exp. Psychol.*, 1940, 26, 568-587.
6. Irmina, M. The effects of summer vacation upon the retention of the elementary school subjects. *Cath. Univ. Amer. Educ. Res. Bull.*, 1928, 3.
7. James, W. Talks to teachers on psychology. New York: Henry Holt and Company, 1939 (reprint), 89-90.
8. Jenkins, J. G., & Dallenbach, K. M. Obliviscence during sleep and waking. *Amer. J. Psychol.*, 1924, 35, 605-612.
9. Jones, H. E. Experimental studies of college teaching: the effect of examination on permanence of learning. *Arch. Psychol.*, New York, 1923, 10, No. 68.
10. Krueger, W. C. F. The effect of overlearning on retention. *J. exp. Psychol.*, 1929, 12, 71-78.
11. Ladd, G. T., & Woodworth, R. S. *Elements of physiological psychology*. New York: Charles Scribner's Sons, 1911.
12. Luh, C. W. The conditions of retention. *Psychol. Monog.*, 1922, 31, 1-87.
13. McGeoch, J. A. *The psychology of human learning*. New York: Longmans, Green & Company, 1942.
14. McGeoch, J. A., & McDonald, W. T. Meaningful relation and retroactive inhibition. *Amer. J. Psychol.*, 1931, 43, 579-588.
15. Pressey, S. L. *Psychology and the new education*. New York: Harper & Brothers, 1933, chapter 11.
16. Pyle, W. H., & Snyder, J. C. The most economical unit for committing to memory. *J. educ. Psychol.*, 1911, 2, 133-142.
17. Radosavljevich, P. R. *Das Behalten und Vergessen bei Kindern und Erwachsenen nach experimentellen Untersuchungen*. Leipzig: Otto Nemnich, 1907. (*Pädagogische Monographien*.)
18. Robinson, E. S. The relative efficiencies of distributed and concentrated study in memorizing. *J. exp. Psychol.*, 1921, 4, 327-343.

19. Steffens, L. Experimentelle Beiträge zur Lehre vom ökonomischen Lernen. *Z. Psychol.*, 1900, 22, 321-382.
20. Strong, E. K. Effect of the time interval upon recognitive memory. *Psychol. Rev.*, 1913, 20, 339-372.
21. Titchener, E. B. A text-book of psychology. New York: The Macmillan Company, 1928, 380-381.

WATSON'S STUDIES OF THE BEHAVIOR OF THE HUMAN INFANT

1. BEGINNING OF EXPERIMENTAL STUDY OF THE CHILD

PRIOR TO the studies of John B. Watson and his co-workers from 1917 to 1920, little actual experimental work had been done upon the growth and behavior of the human infant. There were, to be sure, several excellent biographies of individual children, notably those by Darwin and Miss Shinn. But owing to the resistance and objections of many people, parents in particular, to "experiments" upon the human young, few systematically controlled observations had been made on the appearance and development of behavior patterns in children from which reliable conclusions could be drawn. Watson's pioneer studies, carried out at the Harriet Lane Hospital in Baltimore, did much to remedy the neglect of this field and to emphasize the need and importance of further work. Much of the interest in the study of the young child which we see exhibited today in the establishment of institutes of child welfare and nursery schools had its origin in these early experiments.

While the value of his work in child psychology is fully recognized, Watson is best known not as a child psychologist but as the founder and champion of an objective "school" of psychology known as behaviorism. John B. Watson received the Ph.D. degree from the University of Chicago in 1903 with a dissertation in the field of animal psychology. The next year he went to the Johns Hopkins University, where in 1908, at the age of thirty years, he became professor of experimental psychology. To Johns Hopkins, Watson brought a strong interest in animal experimentation and a growing belief in the practicability of extending to man the objective methods used with so much success in animal work. One of his primary objects, in fact, in initiating the experiments with children described in this chapter, was to demonstrate the value of objective method in genetic studies of the human young.

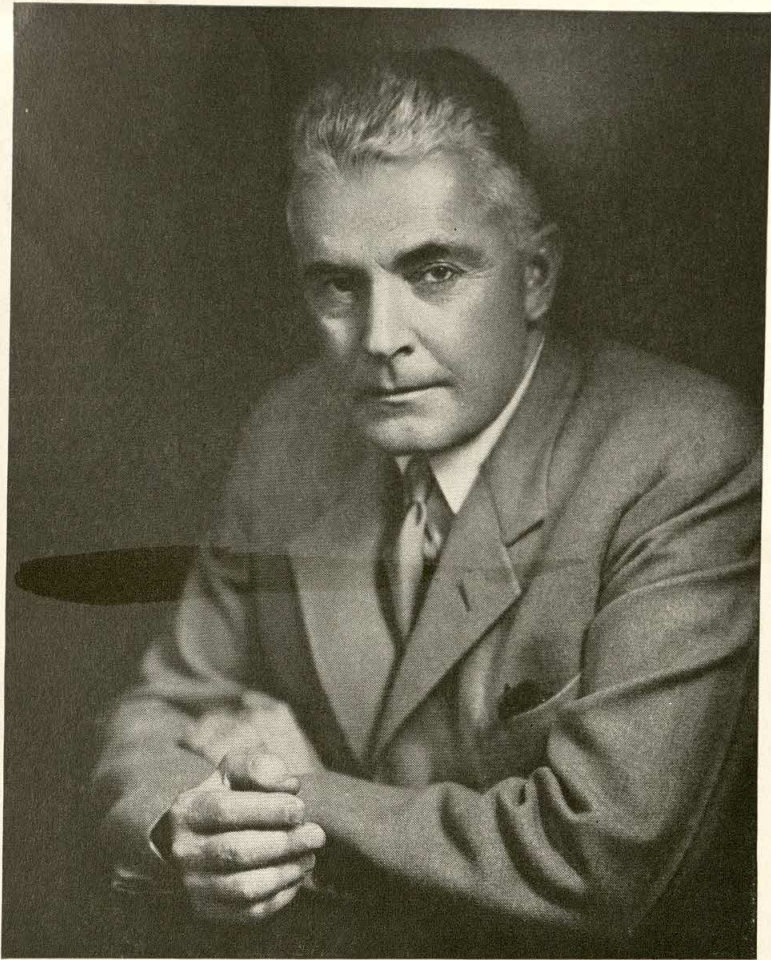
Watson's procedure was to observe infant activity progressively from birth onward; to catalog fundamental reflexive and instinctive tendencies as they appeared together with unlearned emotional expressions; and to

discover how early habits develop from "innate" modes of behavior. Careful note was made of the times at which children begin to reach for, manipulate, and handle objects; sit alone, crawl, and walk. What objects and situations babies are afraid of, angered by, and delighted with were also listed, and the genesis of these emotions studied. Information was obtained, not simply by observing behavior, but by arranging situations and trying them out on children. Such records give a picture of the developing human infant, and enable us to detect and remedy bad habits and perverse emotional attachments at an early age. They have furnished exact data on the old question of whether a person *inherits* specific fears of animals and objects, or whether such fears are learned. Finally, studies like these aid the psychologist in guiding and directing older children toward the kind of education and the kind of vocation for which their talents fit them.

2. MOTOR AND SENSORY DEVELOPMENT

The earliest reflex noted by Blanton, who worked under Watson's direction, was sneezing, which may occur even before the so-called birth-cry. Hiccoughing, yawning, and, of course, crying, also appear very shortly after birth. Besides the birth-cry which is apparently reflexive and comes with the first gasping intake of air by the new born infant, Mrs. Blanton differentiated crying due (1) to hunger, (2) to injury and rough treatment, and (3) to fatigue. It is interesting to note that the regular flow of tears does not ordinarily appear until the child is about a month old. Sucking which includes tongue, lip, and cheek movements, and is followed by swallowing, can be demonstrated in the first hour after birth. The reflexes of elimination occur shortly after birth, and sometimes even before birth. Smiling, which comes somewhat later—usually not before the fourth or fifth day—is also unlearned behavior. If the baby is comfortable and well fed, smiling may be elicited by light touches on the body, by stroking the skin, or by a gentle rocking. Apparently the child does not *learn* to smile voluntarily at some person or some thing until it is at least one month old. Jones in a study of 365 babies, found that conditioned or learned smiling, i.e., smiling when the experimenter smiles or talks baby-talk, rarely appears before the child is 30 or 40 days old. Gesell reports smiling in response to social stimulation at about 90 days.

Careful observation of the new born infant shows that its first activities are in the nature of "mass movements," that is, general diffuse bodily shiftings rather than local movements of separate parts—arms and legs. Apparently almost any stimulus will lead to almost any response, although the reaction is strongest at the point of stimulation, spreading out with



Bachrach

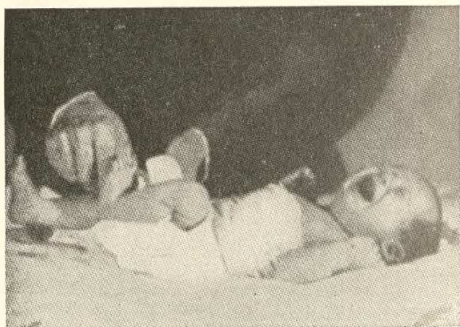
JOHN B. WATSON

1878-



A

A is the grasping reflex of a child twelve days old. B shows the defensive reflex of left foot to slight pinch on inner surface of right knee.



B



C

C is stimulation for the Babinski reflex. The blunt end of a match is rubbed across the sole of the foot. The result is shown in D. The great toe shows extension, the small toes "fanning or flexion." (This is a very variable reflex so far as pattern is concerned.)



D

FIG. 25

SOME INFANT REFLEXES

(From Watson, *Psychology from the Standpoint of a Behaviorist*, 2nd ed. J. B. Lippincott Company, 1924.)

decreasing intensity to other body parts. Pinching the toe, for example, causes general body activity; a bright light provokes thrashing movements of the arms and legs as well as eye movements. Local differentiated movements appear as the child grows older. One of the earliest voluntary motor activities is that of grasping (Fig. 25, A). A baby's fingers will readily close upon a stick, pencil, finger, or any small object. Clinging to this like a small monkey, the child can easily be lifted. All but about 2 per cent of normal children can support their own weight in this way, the time during which they hang suspended varying from a fraction of a second to as long as a minute. This activity is apparently unlearned. It appears shortly after birth and drops out when the child is about 4 months old, although it sometimes persists for a longer time. Once having disappeared, it never returns. The late dropping off of this reflex may be an indication of retarded growth, but at just what age the line should be drawn has not as yet been definitely determined.

If an experimenter stimulates a new born child by pressing with his finger upon its chin, the coördination of the two arms into a single defense movement is at first diffuse and poor. After 4 or 5 days, however, co-ordinated movements of the two arms appear with fair regularity. Figure 25, B shows the "defense reflex" of a baby to a slight pinch on the inner surface of one knee. Note that the opposite leg is drawn up as though to push away the annoying stimulus.

With the waning of the grasping reflex (at about 5 months) the child begins to reach; this activity involves extending the hands for an object, seizing it, and (usually) carrying it to the mouth. Somewhat later (6 months) the baby begins actively to employ its thumb, which hitherto has been well-nigh useless, in opposition to its other four fingers. Reaching activity is tested for by holding some attractive object, say a stick of candy, directly in front of the child and allowing him to try and get it. If he grasps it, he is allowed to taste or suck it. A 6-months-old baby will reach not only for candy but, in general, for any small object held close by. If the object reached for hurts or gives pain, the child soon learns not to reach for that particular thing. After several scorches from a candle, for example, the baby, even at an early age, learns to let lighted candles alone when they come within reach. A somewhat surprising discovery growing out of the study of reaching is that the child does not reach for objects more than two feet away. This is contrary to the poetic notion that the young child will reach for the moon or for any distant object.

Grasping of objects depends upon hand-eye coördinations; and hence it is important to know when eye movements become coördinated and which movements are first to appear. One method of studying eye movements is to place the baby on its back in a dark room with the head held

gently in place by an assistant. Just above the baby's head a perimeter is then placed. This contrivance, which looks like the half of a hoop, carries a small light. The baby's head being the center of the circle of which the perimeter is the semi-circumference, the light when moved is always the same distance from the baby's eyes. By moving the light to the right or left, movement of the child's eyes in following the light can be studied. Watson found that the "roving" movements of the eyes are poorly coördinated at birth. Sherman, who has verified these observations on ninety-six babies ranging from 1 hour to 12 days old, adds that after 30 to 40 hours coördination is fairly good. The first coördinated following movements with true fixation are those in which the baby's eyes turn to the left or right. Somewhat later (after 50 to 80 days) upward and downward movements appear; and the baby when 2 to 3 months old is able to follow the light if it is moved in a circle. At about the same age, the child responds more often to colored than to uncolored objects; and by the age of 6 to 9 months color preferences become marked. Red is the preferred color, followed by yellow and blue. Individual differences in preference for colors appear clearly after about 2 years. The motor response of blinking, brought about by passing the hand between the child's face and a source of light, occurs at 2 to 3 months. This reflex is extremely useful in protecting the eye from injury, and belongs to the general class of movements known as "protective reflexes."

Tickle the sole of a baby's foot or stroke it gently with the blunt point of a pencil and the big toe will turn upward in a fanning or extension movement, while the other toes turn down (flexion movement). This reaction, which is entirely unlearned, is called the Babinski reflex, and is found in practically all infants (Fig. 25, C and D). Sherman reports that in 90 per cent of the cases examined by him, extension at the first stimulus was followed quickly by flexion at the next stimulus. He regards the failure of the second movement to appear as evidence of poor functional or neural development. The Babinski reflex usually disappears between the sixth and twelfth months of the child's life. It is of considerable clinical value, since its presence indicates incomplete development of the nervous system. If the reflex persists longer than one year the supposition is that the baby is not developing normally. When found in adults the Babinski reflex is a sign of a pathological or diseased condition.

An important sign of development is the child's ability to sit alone, for, like reaching, it shows that the baby is gradually learning how to use its own body. At 3½ months, says Watson, the infant may sit unsupported for as long as two minutes. M. C. Jones, on the basis of more observations than Watson, puts the first appearance of sitting alone at

5 months, Gesell at 6 to 8 months. The discrepancy here apparently results from the setting up of different criteria for "sitting alone." By the age of 6 months the child will sit alone twice as long as at 3 months, and will play with its toes, pull at its clothing, and strike the bed upon which it is sitting. After sitting alone, the first step in the locomotion series is crawling, from which standing and walking develop. Many infants can support themselves in the standing position, by holding on to some object, as early as the eighth or ninth month. The age at which a child begins to walk depends upon its health, its weight, and whether it has had any frights or injuries from falls. Walking may appear at 1 year of age or even earlier. After the first step has been taken, actual skill in learning to walk is contingent upon the increase in body strength and growth, as well as upon the praise and encouragement received from the parents. As the sphere of his interests enlarges, the normal child seeks voluntarily to increase his means of getting what he wants.

The following quotation summarizes growth changes over the first five years of life:

In the first year, the child gains control of his basic muscles so as to be able to grasp objects and stand erect and to do many other acts in his new nonaquatic environment. In his second year, typically, he walks and runs and articulates words and phrases. At this time he may acquire bladder and bowel control. In his third year, he speaks sentences and begins to comply with the demands of his home. In his fourth year, he asks many questions and becomes self-dependent in the routines of the house. At five he hops and skips, can tell a long story and is "a self assured conforming citizen of his small world."—Carmichael, *ref.* 4.

3. MENTAL AND SOCIAL DEVELOPMENT

Curves showing changes in intelligence with age will be found in Figure 26, page 132. Nearly all tests of mental ability exhibit the general trend shown by measures of intelligence: rapid upward movement in the early years becoming slower and slower during the teens and later on. By 20 or so the speed and accuracy with which we do mental tasks has about reached its peak—with considerable individual differences. Figure 26 shows the age-progress curves for boys 9 to 14 years old on five mental tests, cart construction, narrative pictures, problem box, object memory, and substitute learning. Cart construction requires that pieces be put together to make a vehicle; narrative pictures that separate pictures be arranged in order to tell a story which the child has just heard; the problem box that levers be successfully manipulated to solve a puzzle. Object memory measures recall memory for twelve small objects shown and named only once; and substitution learning the acquir-

ing of a code whereby nine symbols can be marked each with its appropriate digit. Performance at each age level is expressed as a percentage of the 9-year record. It is evident that all of the tests but object memory showed a marked increase in score with age.

As growth in mental ability proceeds, the young child exhibits an increasing degree of overt emotional control as well as increasing social participation. Temper tantrums (kicking, biting, screaming) reach their peak between the second and fourth years, after which they gradually decline to be replaced by scolding, sulkiness, fussiness, and negativism. The child of 3 to 4 who is non-coöperative and individualistic will by 5 or 6 begin to share his belongings with others.

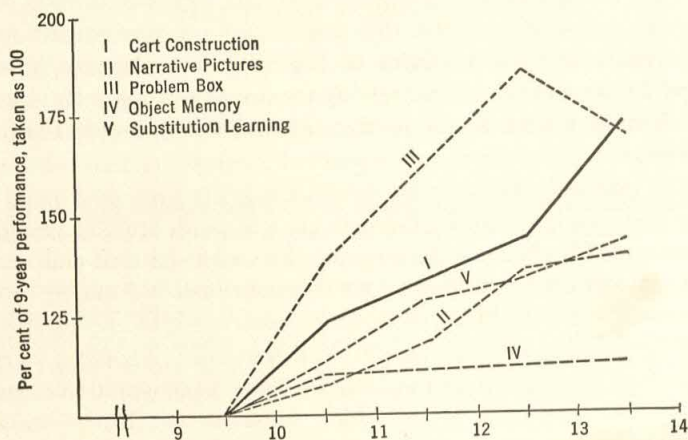


FIG. 26

Performance at each age-level is expressed as a percentage of age 9.5.

The relative contributions to child development of maturation and of learning have been studied through what has been called the method of co-twin control. One member of a pair of twins is given intensive training and his performance later on is compared with that of his untrained twin. In a well-known experiment of this sort, McGraw studied a pair of male twins (Jimmy and Johnny) from birth to about 2 years. Johnny, the trained twin, was given intensive practice while Jimmy, the control, received no coaching. Specific training had little effect upon such fundamental activities as crawling, walking, sitting, reaching, and upon the simple reflexes; but in more complex acts such as diving, swimming, climbing, manipulating boxes, and the like, Johnny far surpassed Jimmy. A certain maturation level appears to be crucial for the appearance of the simple functions. Maturation is necessary also for the more complex be-

haviors but these are influenced to a much greater degree by special training.

When intensive coaching was discontinued, untrained Jimmy gradually caught up with trained Johnny. Intensive practice has little effect until the child's nervous system has reached a developmental level sufficient to "carry" the activity. Teaching is unable, therefore, to effect permanent changes in behavior in advance of a child's maturation level. If training precedes development, time and energy are wasted and the child's interest is lost. Gesell and his associates have confirmed McGraw's findings essentially in several careful co-twin studies.

4. HANDEDNESS IN CHILDREN

As is well known, most adults are right-handed, although a minority (perhaps 10 per cent) use the left as the preferred hand. Is preference for the right hand innate or is it a learned response? The answer to this question is of more than theoretical interest. If handedness is in-born, it is probably safer to let a child use his preferred hand rather than try and change him, as there is clinical evidence that stammering and emotional disturbances often occur when a left-handed child is forced to change over to his right hand. But if handedness is a matter of chance or of early habit-formation, then, since most people are right-handed, it would seem wise to teach all children from the beginning to use the right hand in preference to the left. Watson investigated this question of handedness in several ways which illustrate so clearly the experimental techniques used in child study that they will be described rather fully.

1. First, a comparison was made of the length of time during which twenty very young children could hang from a bar, supporting their own weight either by the right hand alone or by the left hand alone. The tests were begun at birth and continued for the first ten days of the child's life. Results showed that children hang, on the average, as long with the left as with the right hand, i.e., no innate preference in this activity, at least, was apparent (see Fig. 25A, for illustration).

2. Next the amount of random slashing movements made by the baby with its right and its left arms was measured. The record of such voluntary activity as appeared was kept by means of a small "work adder," a device which when operated caused a notched wheel to revolve in one direction. This wheel wound up a weight attached by a cord to a drum thus giving a record of work performed. One end of the adder cord was fastened to the child's wrist, the other to the weight. Two

adders were used so that records from both arms could be taken simultaneously. This experiment, which occupied about five minutes, showed that almost identical amounts of work were done by the right and left arms, thus indicating that neither arm (and neither hand) is favored in gross slashing movements of this kind.

3. In the third test, children from 5 to 12 months old were examined to see which hand was used first in reaching for objects. Twenty babies were tested once a week, ten to twenty trials being given to each. Generally a stick of candy or a candle served as the test object. This was brought in slowly toward the baby on a level with its eyes, and exactly in the middle line squarely between the baby's two hands. Usually reaching occurred when the object was about two feet away. A careful record was kept of which hand was used for reaching; or, if both hands were used, as sometimes happened, which touched the object first. No marked preference for either hand appeared in this test; sometimes the child used the right and again the left hand more often.

4. As a final test, measurements were made of the length of the right and left forearms, wrists, palms, and fingers. No significant differences were found between the right and left measurements of 100 babies.

These studies of handedness suggest the absence of inborn preference for either hand. But other studies indicate that hand preference—if not inborn—develops very early. Thus, left-handedness is often marked at 1 year and may be well established before the age of 2 years. The origin of handedness in a given child is not definitely known, and modern educational practice takes a compromise position with regard to treatment. If the child who shows a strong preference for the left hand finds great difficulty in changing, develops speech disorders, or shows signs of irritability, nervous tension, and unhappiness, he is usually left alone. Changing such a child from left to right hand involves an expenditure in nervous energy which does not justify the result. But children who seem able to change from left to right hand without too great effort—and this is true of most children—are encouraged to shift.

Ours is a right-handed culture. The desire of parents that their children shall not be "different" is of major importance in fitting the growing boy and girl into the right-hand mold. At an early age the boy is taught to shake hands, take off his hat, and write with his right hand; later on he finds that tools, baseball gloves, and golf clubs favor the right-hander. Convenience and social pressure are usually strong enough to make all but the most stubborn left-handers conform. These latter should not be forced to change.

5. TESTS OF DEVELOPMENT IN YOUNG CHILDREN

Tests of physiological development and motor control are often used to determine whether a baby is "normal" for his age. Tests for young children are usually grouped into a graded series or scale from which norms or developmental standards may be calculated. A good illustration of such a scale is Kuhlmann's "Tests of Mental Development." Kuhlmann has assembled and standardized eighty-nine tests which cover the age-range from 4 months to 16 years and beyond. The early age-tests represent, for the most part, methods of discovering whether a child possesses the normal reflexes, as well as the motor, sensory, and language activities characteristic of children of his age. By the age of 6 months, according to Kuhlmann's scale, a child should be able to sit without support one-half minute or more; to react by turning his head toward the source of sound when a telegraphic snapper or small bell is sounded; to reach for objects such as a bell, rattle, colored ball, when dangled before his eyes within reach; to pick up a red checker by a pincer-like movement using thumb in opposition to fingers. By the time he is 1 year old, a child should—among other things—be able to remove a cloth cover from a cup or rattle, imitate arm movements, crawl or creep three or more feet to secure an interesting toy, mark with a pencil on paper. At 2 years, the child should obey simple commands such as "Go bring me the ball"; recognize objects such as a block, key, watch, doll; name familiar objects seen in a picture, e.g., apple, flower, bird. In accordance with his successes and failures with these tests, the mental age of the baby can be calculated in much the same way as that of an older child.

Another set of performance tests for the young child is the Merrill-Palmer series of tests drawn up and standardized by Rachel Stutsman. These tests (there are ninety-three in all) are grouped into 6-month intervals from 18 months to 6 years; that is, the first group of eleven tests is for children 18 to 24 months old; the second group (ten tests) for children 24 to 30 months old; and so on. The mental age of the child is calculated from the total number of tests in the series which he is able to pass. To illustrate the kind of performances called for, the average child of 2 to $2\frac{1}{2}$ years is expected to be able to repeat four simple words (e.g., *kitty*, *ball*); to recognize himself in the mirror; fit sixteen cubes into a box in 125 seconds or less; put pegs into a board; pull a stick attached to a string in toward himself, using one or both hands; arrange four cubes—of different sizes all open at one face—into the

form of a "nest," each cube being fitted into the one next larger in size; answer six out of ten simple questions, e.g., "What does a kitty say?" "What is this (chair)?"; cut paper with scissors; and repeat simple word-groups, such as "see the pretty dollie." As is evident, these tests measure the child's ability to recognize and manipulate familiar objects of his environment in a sensible if elementary fashion.

The 1937 Stanford-Binet (p. 224) since it begins at 2 years is often used with young children, as is also the Wechsler Intelligence Scale for Children (p. 230) which begins at 5 years. Various other test series have been developed specifically for use with the preschool child, e.g., the California First Year Mental Scale, the Cattell Intelligence Scale, the Minnesota Preschool Tests, and the Vienna Babytests. The last two of these will be briefly described. The Minnesota scale consists of both language and non-language items; it may be scored to show either verbal or non-verbal development or both (combined scale). Materials employed in the test are interesting to young children and varied enough to cover a wide range of activities. The tests span the age-range from 18 months to 6 years, and must be administered by a trained psychologist. The Vienna Babytests, developed by Charlotte Buhler and her assistants, cover the first and second years of life. Four areas of activity are sampled by these tests: (a) bodily control and coördination; (b) mental ability; (c) object manipulation, memory, and attentiveness; (d) social development, e.g., smiling when another smiles, actively seeking contacts with people, play, use of toys and the like. No claim is made that a given test samples only *one* of the four fields, but the division is convenient and useful in diagnosis.

Tests of very young children do not have very precise predictive value for the school years. This is owing in part to the real difficulties encountered in testing very young children. A baby's attention is easily distracted and difficult to hold once secured. It is hard also to get an adequate measure of the shy, timid child, or of the boisterous, spoiled, and stubborn child. Physical condition, fatigue, emotional upsets all make for variability and serve to decrease the significance of a test rating. Even more than with older children, the personality of the examiner is important; and a person unacquainted with children can rarely get coöperation. Again, tests for very young children are of necessity concerned largely with sensory and sensory-motor functions. Such tasks are not highly correlated with the verbal (abstract) items found in general intelligence tests for older children. In spite of real obstacles to an accurate rating of the very young child, measurements made at later age-levels do not often change the child's *general* classification. Thus while the IQ of a child may drop from 130 to 120 between the ages of two and

eight, or may increase from 70 to 80, the designation of the child as bright or dull remains unchanged.

Seven simple performance tests, much used with younger children, are shown in Figure 27. These tests all measure, in a general way, the child's muscular development and hand-eye coordination. They gauge, too, his perception of form and shape, and knowledge of common everyday things. The Wallin Peg Boards (A and B) set before the child the task of fitting the six round pegs in A, and the six square pegs in B, into the holes in the board. These are very easy tests and can be done by normal children from 18 months to 2 years of age. They are often used as tests of the feeble-minded and the low-grade. The Nest of Cubes Test

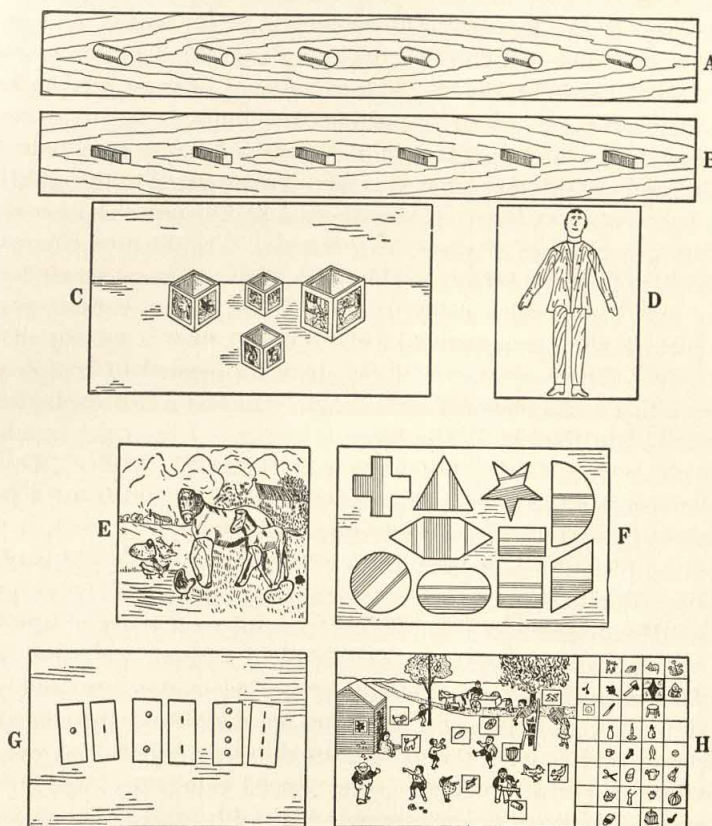


FIG. 27

SOME PERFORMANCE TESTS MUCH USED WITH YOUNG CHILDREN

A and B are Wallin Peg Boards; C is the Nest of Cubes Test; D, Pintner's Manikin Test; E, the Mare and Foal Test; F, the Seguin Form Board; G, the Decroly Button Test; H, the Healy Picture Completion Test, No. 1.

(C) has been described above. Each cube is to be fitted into the one next larger in size. The average child of 2 to 2½ years should do this test in four minutes or less. In *D*, Pintner's Manikin Test, the child is required to fit the arms, legs, and headpieces, which are separate, onto the trunk to form a man. This test is rarely done correctly by children under 8 years of age. The Mare and Foal Test (*E*) sets the task of fitting certain cut-out pieces into their correct locations. The normal 3-year-old takes about five minutes to do this test. The Seguin Form Board Test (*F*) is the best known of the many form-board tests. The average child of 3 years can fit the blocks into their proper places in about two minutes.

The Button Test (*G*) was devised by the Belgian psychologist, Decroly. The six strips are of flannelette, three inches by six inches, and are arranged in pairs. The first pair consists of a strip with a button sewed on it and a second strip in which a button-hole has been worked to fit the button. The next pair of strips contains two buttons and two button-holes, respectively, and the third, four buttons and four button-holes. The buttoning and unbuttoning process is first demonstrated to the child, who is then encouraged to try it. A normal child of 2½ years can usually do one button, a child of 3 years, two buttons. The Picture Completion Test, No. 1, of Healy (*H*) is usable with children from about 5 to 12 years of age. The child is called upon to fill in certain cut-out portions in the picture with appropriate blocks selected from a set supplied by the examiner. There are more blocks than are needed to complete the picture, so that a real choice must be made. The test is first explained and demonstrated to the child. The score is computed from the number of correct placements. Of the tests shown in Figure 27, *A*, *B*, *C*, *D*, *E*, *F*, and *G* are used in the Merrill-Palmer series; *D*, *E*, *F*, and *H* are a part of the Pintner-Paterson Scale of Performance Tests.

Gesell, in his studies of preschool children at the child clinic of Yale University, devised an extensive schedule of tests and observations to be used with children from about 1 month to 6 years of age. Four fields of behavior activities are sampled: (1) *Motor behavior*, which deals with muscular capacity, coördination, locomotion, and the use of hands and arms; (2) *language*, involving vocalization, word comprehension, speech, and conversation; (3) *personal-social behavior* as exhibited in reactions to persons, initiative, play, social experience, and information; and (4) *adaptive behavior*, concerned with hand-eye coördination, imitation, discrimination, ability to make proper adjustments to a given situation, and sensible control. Gesell's schedule is not a test in the sense of being concerned primarily with determining a score or mental age. It is rather a summary of the child's developmental status based upon

physical and psychological examinations as well as upon clinical observation.

6. EMOTIONAL DEVELOPMENT OF THE YOUNG CHILD

As a result of many observations carried out on very young children, Watson reached the conclusion that there are only three fundamental emotions, fear, anger (or rage), and love. These "primary" emotional patterns appear at birth or shortly thereafter. Watson's method of testing for native emotions was to bring the child into the laboratory, present stimuli which are known to produce emotions in adults, and note carefully the child's typical reactions. In every case an effort was made to describe the emotion in terms of its characteristic pattern, and to define it strictly in terms of the situation which called it out.

What stimuli, for example, will call out the emotion of fear in the young child? After working with various situations, Watson discovered that *loud noises*, *painful stimuli*, and the *sudden loss of support* (allowing the child to drop a few inches) cause the infant to recoil, catch its breath, clutch with its hands, and if possible crawl away. These responses constitute the "pattern" of what Watson called a fear-reaction. No other stimuli brought out fear so regularly and definitely; for instance, these young children were found to have no native fear of the dark, of animals, of people, or of the thousand and one things of which children are supposed to be "naturally" afraid.

To bring out the emotional response of rage or extreme anger Watson found it sufficient to hamper or restrict the infant's movements in some way. If the child's head is held lightly between the hands, or the movements of its arms or legs restrained, it will exhibit the behavior which is called rage. This "pattern" consists in a stiffening of the body, crying or screaming, and struggling to escape. If first attempts are unsuccessful, the child will hold its breath or scream until it becomes blue with rage.

In very young children the emotion of love, as Watson defines it, is characterized by smiling, gurgling, and cooing. Sometimes the child moves its arms and legs rapidly, and sometimes it lies very still, the picture of contentment. The stimuli for love responses are a stroking of the skin, tickling, patting and rocking. When the child becomes older, it will hold out its arms to be taken by the adult.

Other investigators have found that Watson's situations do not *always* lead to fear, anger, or affection as he supposed. For example, some young children show no fear of loud noises or of the threatened loss of support, though most children do. Later research, also, has failed to reveal the definite and clear-cut *patterns* of rage, fear, and love which Watson

believed to be the only "primary" emotions. An experiment by Sherman is enlightening on this point. A group of thirty-two graduate students in psychology were shown motion pictures portraying the reactions of infants to four kinds of stimuli—hunger, sudden dropping through a distance of two or three feet toward a table, restraint of head and face, pricking with a needle. The stimuli themselves were not seen by the observer, as the scenes showing the administration of the stimuli had been deleted from the film, leaving only those showing the child's reactions. Little agreement appeared when the student judges recorded what they thought was the emotional reaction exhibited by the child in each instance. Thirteen observers called the hunger response, anger, seven called it hunger, seven fear, three pain, one grief, and one consternation. Fourteen called the response to dropping anger, only five fear. In another test, a group of medical students and nurses were shown the stimuli which produced the emotional responses being administered, as well as the responses themselves. Judgment under these conditions was much improved, fear being most often named for the dropping responses, anger for restraint, and pain for the needle prick. When the films were arranged so that the infant's emotional behavior was shown preceded by some stimulus other than that which actually produced the emotion, accuracy of judgment was much decreased, suggesting that the observers had based their opinions more on the stimulus situations than on the behavior. This finding probably holds true for other than laboratory situations. We judge a person's probable emotional state from the *expected* effect of the stimulus upon him more often than from his overt responses. A given emotion is not revealed by a stereotyped and invariable sequence of movements. Fear may be "typically" characterized by shrinking of some sort, and rage by aggression or attack, but the extent and variety of the behavior differs greatly from child to child and for the same child on two or more occasions.

7. HOW CHILDREN ACQUIRE FEARS AND OTHER EMOTIONAL ATTITUDES

In searching for situations and objects which might lead to emotional behavior, Watson carried out several tests on older children, i.e., those varying in age from about 4 months to 1 year. All of the children examined had been brought up in the hospital and had never seen any of the animals or other objects used in the tests.

First, the child was brought into the laboratory and allowed to sit upon the lap of its mother or an attendant. Various animals were then presented and the baby's reactions noted. When a very lively black

cat purred near by, the baby showed no fear; nor did it show any fear of a pigeon, a rabbit, or a white rat. These were all reached for, one baby trying to put the rabbit's ears into her mouth. In addition to the tests in the laboratory, each child was taken to the zoo and allowed to get quite close to the animals. The result was the same as before: not the slightest fear was shown.

These tests are interesting and to many, perhaps, highly questionable. If originally a baby shows fear only when confronted by a loud noise or when threatened with loss of support or when actually injured, how does it come about that older children and adults are afraid of so many things: the dark, snakes, dogs, bugs, cats, and far more innocuous stimuli? The answer given by Watson is that they have *learned* to fear these things, and that the method by which such fears have been learned is the familiar conditioned response.¹ A conditioned response, as we know, is behavior called out by some stimulus other than that to which it was originally bound. Strike a dog with a stick and yell "Go away," and it is probable that at a later time the dog will not wait to be struck, but will run when he sees the stick, or even when he hears your voice. This is a simple illustration of a conditioned response. Fear and running away are brought out in the dog by the yell, or by sight of the stick stimuli originally powerless to produce these specific reactions. In the same fashion, fear of lightning may be instilled in a young child by the fact that lightning and thunder (loud noise-original stimulus for fear) occur together. Many examples of the same kind are cited elsewhere (see pp. 53-55).

Watson decided to see whether he could build up a conditioned fear in the laboratory. His subject was a boy, Albert B., 11 months old, who was possessed of a stolid and phlegmatic disposition. First it was determined by actual test that Albert was not afraid of furry animals, such as the rabbit and the white rat. When put within reach he immediately attempted to grasp them. However, Albert's reaction to the loud clang made by striking a steel bar with a hammer was distinctly one of fear. He would pucker up his lips, throw up his arms, turn over on his side away from the noise and begin to whimper. The problem which Watson set was this: Can an animal be "substituted" for the loud noise and thus become a fear-object? The experimental setup was as follows. First, the white rat was presented to Albert. At once he reached for it, and as soon as he touched it the steel bar just behind the child's head was struck a heavy blow. The fear reaction immediately appeared. The next time the child reached for the rat the noise was repeated with the same result—fear was distinctly shown. Seven days later the child eyed the rat warily when it was presented, refusing to touch it. When the rat

¹ See Chapter 1.

touched the baby's hand, the hand was quickly withdrawn, but the child did not cry. It was clearly apparent, however, that Albert, while not quite afraid, was not favorably inclined toward the rat. No fear response at all was shown toward other objects, such as building blocks, which had not been presented along with the noise. Three more joint presentations of rat plus noise were now given, and these were sufficient to produce unmistakable signs of fear when the rat was later presented alone. Two more joint stimulations, and the instant the rat appeared on the scene the baby began to cry and to crawl away as fast as he could. Five days later Albert was still afraid of the rat.

In order to discover whether this conditioned fear had spread to other objects, Watson then tested Albert with a rabbit, a dog, a fur coat, and cotton wool. To all of these the child was sharply negative: either he cried or crawled away or both; in every case he refused to play with the objects. None of these things, it must be remembered, had been presented along with the loud noise. Yet because of their similarity, the fear reaction had been generalized to all of them. But no fear was exhibited whatsoever to building blocks, the conditioned fear apparently having spread only to furry objects.

Watson did not attempt to build up any rage reactions experimentally, evidently because of the danger involved in dealing with such behavior; nor did he condition any love reactions in the laboratory. Conditioned or learned rage reactions are often encountered, however, as when a child goes into a temper tantrum or into a screaming fit in order to get candy or some other desired object, or to have its way. We are all familiar, too, with the ease with which a child's love responses can be transferred from mother to nurse or to grandmother, depending upon which one provides food, petting, and comfort.

Research in child psychology has greatly expanded since Watson's pioneer experiments. The present-day worker in the field of child development is essentially interested in the broad aspects of emotional, social, physical, and intellectual growth as these are related to each other and to everyday living. Literally scores of studies have been made of the child in school, his behavior problems, nervous habits, social difficulties, and the like. To illustrate, studies of anger in children from seven months to seven years show this emotion to be typically overt, violent and unrestrained in the very young; personal, relative and socially motivated in the older child. Irrascibility has been linked with poor health, over-indulgence, over-solicitude, and with hypercritical parental attitudes. Apparently, most fears in young children can be tied up with poor training at home or with lack of good training. In many young children (2 to 4 years) sympathetic, friendly and coöperative behavior appears to be a

consistent and characteristic trait. It increases with chronological age, and with mental development; is about equal for boys and girls; and differs greatly from child to child.

8. METHODS OF BREAKING UNDESIRABLE HABITS

Experiments like those described in the last Section are valuable in showing how useless and even harmful emotional responses may be called out by essentially innocent situations. Discovery of the causes which lie behind unwanted behavior and the formulation of methods of treating and eradicating such responses is a major problem for the child psychologist. If, working backward, we are able to uncover the reasons why a child is afraid of a dog or a dark hallway, or why he is thrown into a tantrum by a seemingly innocent circumstance, we are certainly in a better position to clear the matter up. Much unhappiness later could undoubtedly be prevented if harmful habits could be broken up or "re-conditioned" in childhood.

Several experimental studies of the various methods of eliminating harmful habits have been undertaken, among the most interesting of which is an investigation of unnatural fears in childhood carried out by Mary Cover Jones. About seventy children in all, from 3 months to 7 years old, were the subjects. These children were being maintained in an institution for the temporary care of children; they were from fairly good homes and were of normal intellect as determined by individual mental examinations. Children who showed marked fear of snakes, rats, rabbits, frogs, "scare" faces (masks), loud sounds, and so on were selected for study. The primary object was the removal of these useless and hampering forms of behavior. The following paragraphs will serve to illustrate the procedures employed.

1. *The Method of Elimination through Disuse.* This technique is grounded in the common notion that a childish fear will gradually wear itself out (be extinguished, p. 8), if the child is shielded from the fear-inspiring object. This may be true when the interval extends over long periods of time during which the child grows in self assurance and enlarges his experiences, but it did not work very well in Jones's test, which extended over several months. At best, ignoring a fear, or shielding the child from fear-objects would seem to be a temporary expedient. Fears cannot all be so treated, nor can the child be shielded all of its life.

Disuse is often accomplished by means of change-of-environment—a method much used in therapy. Thus, an overworked and harried person may be advised to take a trip, the supposition being that upon return the irritating stimuli will be weakened. Change-of-environment is most use-

ful, perhaps, when combined with other devices which seek to change the person's attitude.

2. *The Method of Verbal Appeal.* Many parents assume that a child can be "reasoned with" or talked out of his fears, of say, snakes or dogs; that by telling stories woven around the fear-object, showing pictures, and the like, the child's curiosity and interest will be sufficiently aroused to overcome his fear. This plan was tried with a 5-year-old girl who exhibited marked fear of a rabbit. Picture-books of "Peter Rabbit," toy rabbits, and stories were used to create interest in "real" rabbits. At the end of a week of such treatment, however, the child's fear of the real rabbit was as strong as before. Apparently verbal assurances are not sufficient defense. When a strong fear urge is present, actions not only speak louder than words, they speak more truthfully.

3. *The Method of Negative Adaptation or Toleration.* Negative adaptation in plain language means that familiarity breeds contempt or at least indifference. The assumption is that the child will finally become "used to" snakes, bugs, or even "ghosts" in a dark room, if such situations are encountered often enough. This belief is fairly well founded. In several instances a child, originally very much frightened by a white rat, became at least tolerant, if not exactly friendly, when the animal was often seen around. This method would seem to be valuable when used carefully, and if the stimulations are not too frequent nor too drastic.

Another form of the negative adaptation method is the "exhaustion method." A child will often whine and tease for some favor as long as there is hope of eventual success. If ignored long enough he may finally give up and resort to a more socially acceptable approach. In like manner, an animal on first being confined in a cage will actively and persistently seek to escape. Failing in this, he finally "gives in" and after a time accepts captivity with an appearance of equanimity. In the case of children, the exhaustion technique is usually more trying upon the experimenter's patience than upon the subject's emotional stability. Swift and immediate punishment is often more effective in ending undesirable behavior than is the "wearing-out" process; and in many cases is more justifiable.

4. *The Method of Repression.* Children's fears are often repressed or temporarily hidden under the surface when they are ridiculed by other children or by adults. Little boys of around four or five will often make a brave show, their desire not to be eclipsed or made to feel inferior serving for the time to overcome their real feelings. But the repression of a fear in this way is dangerous business, as the remedy may turn out to be worse than the original ailment. Instead of being lost, a fear may be the more strongly entrenched as a result of social ridicule and teasing.

5. *The Method of Distraction.* This technique, much used by mothers in everyday life, consists in presenting an object or other stimulus so pronouncedly positive in appeal that the child's attention is temporarily, at least, distracted from the fear-object. A child's fear of a frog or rabbit can be lessened by placing toys and other desirable playthings near the animals. Verbal distractions, too, such as calling the animal by name, pointing out characteristics of the "bunny" or "hop-toad" are useful in inducing a child to forget his fear in the new interest. But the presence of an adult (who does the distracting) makes it impossible to tell how much of the child's new confidence is due to the reliance and trust which is provided by the presence of the grown-up. In many cases, however, the fear may be temporarily lightened through distraction if not completely overcome.

6. *The Method of Positive Conditioning.* Here an attempt is made to associate with the fear-object some stimulus known to be capable of calling out a strong positive (pleasant) response. The hope is that in time the fear-object will itself bring out a positive rather than a negative reaction (p. 142). In several experiments, food was used as the positive stimulus-object to overcome a specific fear of animals, e.g., a rabbit. The procedure was to put the rabbit upon the table on which the child's food had been placed, at a sufficient distance not to interfere with the child's eating, but close enough to keep the animal well in sight. At first the child ate with a wary eye on the rabbit; and for several days this continued with nothing happening. Gradually the animal was moved closer and closer to the food until finally tolerance changed to indifference and even to the positive response of reaching and stroking the rabbit. While successful in many cases, this technique requires careful handling. If fear of the animal or other object is very intense, it may produce a negative response to the food, or such violent crying and screaming as to make the child ill. But if care and patience are exercised, positive conditioning is one of the most effective methods of removing useless and hampering fears.

7. *The Method of Social Stimulation.* There are many cases in which this approach may be used successfully. When a child discovers that other children—as well as grown-ups—are not frightened at the sight of dogs, rabbits, bugs, and so on, curiosity and self-assertion are often powerful enough to "overcome" his fear without more ado. To some extent this is true also of adults, unless the fear is deeply ingrained as the result of intense experience; or else serves some ulterior purpose. Girls, for instance, are no more afraid of small animals and bugs than boys up to the period of adolescence. When the "grown-up" period is reached, girls find that it is considered "feminine" and rather commendable by the

young males if they scream and show fright at such objects. Hence a fear, partly feigned, is developed. If an adult is in earnest about ridding himself of a foolish and useless fear, the original cause must first be found and a period of reëducation instituted until the fear-stimulus becomes no longer effective.

Of the seven methods outlined above, perhaps the most practical are conditioning and social stimulation. While sometimes effective alone, verbal appeal, elimination through disguise, negative adaptation, repression, and distraction are most valuable when employed together with some other technique. Which method to employ will depend upon the kind of fear and its intensity, as well as upon the child's general health, intelligence, and environmental background. Work on problems like these is going on in many laboratories through the country. Watson's pioneer studies marked the beginning of the widespread activity and interest evinced today in the psychology of the young child.

BIBLIOGRAPHY

1. Bayley, N. The California first year mental scale, Univ. of California Syllabus Series, 243. Berkeley: Univ. of California Press, 1933.
2. Blanton, M. G. The behavior of the human infant during the first thirty days of life. *Psychol. Rev.*, 1917, 24, 456-483.
3. Buhler, C. The first year of life. New York: The John Day Company, 1930.
4. Carmichael, L. Chapter 4, p. 78 in *Foundations of psychology*, by Boring, Langfeld & Weld. New York: John Wiley & Sons, Inc., 1948.
5. Cattell, P. The measurement of intelligence of infants and young children. New York: Psychological Corporation, 1940.
6. Darwin, C. A biographical sketch of an infant. *Mind*, 1877, 2, 285-294.
7. Furfey, P. H., & Muehlenbein, J. The validity of infant intelligence tests. *J. genet. Psychol.*, 1932, 40, 219-223.
8. Gesell, A. The mental growth of the pre-school child. New York: The Macmillan Company, 1928.
9. Gesell, A. *Infancy and human growth*. New York: The Macmillan Company, 1928.
10. Gesell, A., & Thompson, H. Twins T and C from infancy to adolescence: a biogenetic study of individual differences by the method of co-twin control. *Genet. Psychol. Monog.*, 1941, 24, 3-122.
11. Goodenough, F. L. *Anger in young children*. Minneapolis: University of Minnesota Press, 1931.
12. Goodenough, F. L., Foster, J. C., & Van Wagenen, M. J. *Minnesota pre-school tests*. Minneapolis: Educational Test Bureau, 1932.
13. Guthrie, E. R. *The psychology of learning*. New York: Harper & Brothers, 1935, chapters 5 and 9.
14. Jersild, A. T. *Child psychology*. New York: Prentice-Hall, Inc., 1936, chapter 5.

15. Jones, M. C. The elimination of children's fears. *J. exp. Psychol.*, 1924, 7, 382-390.
16. Jones, M. C. The development of early behavior patterns in young children. *J. genet. Psychol.*, 1926, 33, 537-585.
17. Kuhlmann, F. Tests of mental development. Minneapolis: Educational Test Bureau, 1939.
18. Lederer, R. K. An exploratory investigation of handed status in the first two years of life. *University of Iowa Studies in Child Welfare*, 1939, 16, No. 2.
19. McGraw, M. Growth: a study of Johnny and Jimmy. New York: D. Appleton-Century Company, Inc., 1935.
20. Murphy, L. B. Social behavior and child personality. New York: Columbia University Press, 1937.
21. Pintner, R., & Paterson, D. G. A scale of performance tests. New York: D. Appleton-Century Company, Inc., 1917.
22. Pratt, K. C. Nelson, A. K., & Sun, K. H. The behavior of the newborn infant. *Ohio State University Series: Contributions in Psychology*, 1930, No. 10.
23. Sherman, M. The differentiation of emotional responses in infants. *J. comp. Psychol.*, 1927, 7, 265-284.
24. Sherman, M. & I. C. Sensori-motor responses in infants. *J. comp. Psychol.*, 1925, 5, 53-68.
25. Shinn, M. W. The biography of a baby. Boston: Houghton Mifflin Company, 1900.
26. Staples, R. The responses of infants to color. *J. exp. Psychol.*, 1932, 15, 119-141.
27. Stutsman, R. Mental measurement of pre-school children. Yonkers, New York: World Book Company, 1931.
28. Travis, L. E. Speech pathology. In Murchison, C. (Editor), *Handbook of child psychology*. Worcester, Massachusetts: Clark University Press, 1933, 650-698 (see esp. pp. 689-695).
29. Watson, J. B., & Morgan, J. J. B. Emotional reactions and psychological experimentation. *Amer. J. Psychol.*, 1917, 28, 163-174.
30. Watson, J. B., & Rayner, R. Conditioned emotional reactions. *J. exp. Psychol.*, 1920, 3, 1-14.
31. Whipple, G. M. (Ed.) *Child development and the curriculum*. Thirty-eighth Yearbook of the National Society for the Study of Education, 1939, Part I.
32. Wile, I. S. Handedness, right and left. Boston: Lothrop, Lee & Shepard Co., 1934.

8

CANNON AND EXPERIMENTAL STUDIES OF THE EMOTIONS

1. PSYCHOLOGICAL THEORIES OF THE EMOTIONS

THE EXPERIMENTAL study of the emotions offers some of the most fascinating and at the same time most elusive problems for research in psychology. Reasons for this are not hard to find. Emotions tend to dissipate when examined introspectively and to lose color and intensity when recalled. It is not easy to get genuine emotional states under laboratory conditions; and numerous technical difficulties arise when the attempt is made to measure complex physiological changes in order to square these off against the subject's verbal report of his felt emotion. However, such obstacles as these have acted as a spur rather than a check to experimental investigation.

The importance of the emotions in behavior has long been recognized by psychologists. But it is only recently that they have come to realize keenly that the success or failure of an individual in his occupation, as well as in his private life, is contingent upon emotional balance as much as, if not more than, upon intellectual ability. The stimulus to investigation, therefore, has come from practical as well as from theoretical considerations.

What do we mean by an emotional response? How does it arise? Upon what factors does it depend? Historically, nearly all writers on the subject have in some way connected emotions with the viscera—the heart, stomach, lungs, and other internal organs. The ancient Greek philosophers located the “seat” of the emotions in the heart or abdomen. Such expressions as “hard-hearted,” “spleeny,” “to eat one’s heart out,” “bowels of mercy,” show the influence of venerable notions which have survived in the common speech of our times.

Several views of the emotions seem to be fairly well defined today, despite the existence of variants and modifications. The *first* of these—the “common-sense view”—regards emotions as primarily mental or conscious phenomena, followed, to be sure, by organic and circulatory

changes, but not dependent upon these. The difference between a man running a race and one running away from a mad dog lies, according to this view, not in the physiological activities of the two men, which are somewhat alike, but rather in their very diverse mental states. A *second* view considers an emotion to be the subject's conscious awareness of the organic and physiological processes going on in his own body. This is the famous James-Lange theory of the emotions, proposed by William James in 1884 and independently by the Danish physiologist C. Lange in 1885. In the average man's view, a person cries when he is sorry and laughs when he is happy or joyful: the bodily expression *follows* the sorrow or joy. Commenting upon this common-sense notion, James wrote: "My thesis on the contrary, is that the bodily changes follow directly the *Perception* of the exciting fact, and that our feeling of the same changes as they occur *IS* the emotion." "We feel sorry," he says, "because we cry, angry because we strike, afraid because we tremble." Partly because of its novelty, and partly because of the authority of its authors, this view of the emotions has exerted much influence upon the thinking of later workers.

A *third* view stresses the part played in emotions by visceral and skeletal muscles as well as by glandular changes, but places considerable emphasis, too, upon the conscious processes involved. Woodworth and Cannon represent this position. An emotion, according to Woodworth, is a "stirred up state of the organism" in which there is an impulse or conscious attitude tending toward some definite activity. It should be noted that the term "stirred up" means active or agitated—not chaotic. First comes the stimulus; then the tendency, impulse or motor set, e.g., to escape or fight, which in turn is followed by physiological changes. The organic components are biologically valuable in that they prepare the animal or man the better to run or fight, as the case may be. Feats of strength and endurance—the run from Marathon to Athens, unbelievable deeds on the football field, heroic actions in the face of danger—all illustrate the value of an emotion in reinforcing the dominant drive and giving impetus to behavior. The factor which, according to this view, marks one emotion off from another is the *set* or *impulse*, which is, in turn, governed by the total situation. Typically, then, emotion *follows* impulse. A man sees a car bearing down on him; he jumps back hastily to the curb, fear surging over him *as* or *after* he jumps. The behavioristic approach represents still a *fourth* view. Consciousness is completely ignored. We have done all we can scientifically, according to the behaviorist, when we have described the motor and glandular reaction patterns which accompany what we call an emotion. The difference between anger and joy is simply the difference in the way the body re-

acts. The individual's mental state is of no consequence, at least to the extreme behaviorist, though it may be the basis for a verbal report of how he feels.

2. PROBLEMS AND METHODS IN STUDYING EMOTIONS

From these different, and to some degree contradictory, points of view, various problems involving the physical basis, as well as the isolation and measurement of the emotions have been attacked by psychologists. The following problems are fairly typical of the experimental work in this field. Although by no means exhaustive, they will help to center the discussion and make it more concrete.¹ (1) First are those studies which deal with the *physical basis* of emotions and the *physiological changes* in emotional states. What rôle do the brain and the nervous system play in emotion? Can an emotion exist apart from the bodily changes—flushing of the skin, rapid heartbeat and breathing, clenched fists, and so on—usually associated with it? (2) In another group are experiments which have sought to discover whether emotions can be *measured* by instruments; whether there are any fairly *definite* and *recognizable* emotional patterns; and whether emotions can be differentiated on the basis of the physiological changes corresponding to the subjective emotional states reported by the subject. (3) Still other experiments have been planned to discover whether emotional states are highly correlated, i.e., whether a person who reacts in a highly emotional way to one stimulus will tend to react with equal intensity to other stimuli. (4) Finally, there are the attempts to reach and resolve emotional difficulties (complexes) by ferreting out the cause of the original upset in the subject's history. This list will serve at least to illustrate how wide and varied has been the experimental attack upon the subject.

In general, two methods have been employed by psychologists in investigating emotions: *the method of expression* and *the method of impression*. The method of expression is the more objective. What it seeks to do is to record and measure the bodily changes which accompany emotional states, putting little or no emphasis upon introspective reports. The method of impression, on the other hand, studies emotions introspectively or, more exactly, retrospectively. The subject, for instance, tells which of two color combinations he prefers, or which of two musical selections has moved him most. Again, after an emotional experience, the subject may be asked to report upon his thoughts and feelings. Were there conflicts, feelings of embarrassment, or anger? When and how did

¹ A discussion of the development of the emotions in children will be found in Chapter 7.



Bettmann Archive

WALTER B. CANNON

1871-1945

they appear? Sometimes the emphasis is on one method and sometimes on the other, but most experimenters have combined the two, taking objective records wherever possible, and recording also the verbal reports or introspections of the subject.

3. THE PHYSICAL BASIS OF THE EMOTIONS

The question of the extent to which an emotion depends upon its accompanying bodily changes is largely a matter of physiological research, and in consequence experiments in this field have been performed for the most part by physiologists. One notable experiment is that of Sherrington, the eminent English physiologist. Sherrington transected the spinal cords of a number of dogs in the neck region, thereby severing all connection between the brain and the body organs of the trunk. All sensations from the viscera and the skin were completely destroyed by this operation, only the sensations arising from stimulation of face, head, and neck region remaining intact. In spite of this drastic reduction of the animal's source of sensory impressions from its own body and skin, emotions persisted with little or no apparent reduction in intensity. One operated dog, for example, showed marked rage and antipathy toward an attendant. She growled, snarled, and bared her teeth, giving the same picture of anger seen in normal animals. At the same time this animal displayed affectionate behavior toward the person who customarily fed and stroked her. We do not know, of course, whether this animal actually *experienced* an emotion. But these experiments certainly prove that emotional expression *can* take place in the absence of the usual visceral changes.

An interesting feature of Sherrington's experiment was the retention by a dog of what might be called "disgust" behavior. Normal dogs, Sherrington found, refuse to eat dog meat, no matter how much it is disguised by being mixed with other foods or covered with milk. Always they exhibit disgust and aversion toward it; probably the odor has much to do with this. Sherrington's "spinal dogs," as these operated animals were called, exhibited the same disgust when given a bowl containing dog meat, refusing to touch it or eat it no matter how much they were urged.

Significant work on the nervous and physiological basis of the emotions has been carried on by W. B. Cannon of the Harvard Medical School, and his associates. Cannon investigated with much care the rôle of the autonomic nervous system in emotional activity. The autonomic nervous system, as its name implies, is normally active in those functions of the body which involve little or no volition, or knowledge, on the part of the subject. The autonomic nervous system controls the vegetative

functions of the body—the action of the heart, lungs, stomach, and internal glands, as well as circulatory and sweat-gland changes. The two main branches of the autonomic are the thoracico-lumbar or sympathetic (the middle part) and the cranio-sacral (the upper and lower parts). Cannon found that the sympathetic branch is active in intense or unpleasant emotions such as fear and rage. Its general function in the body is to intensify certain sorts of activity; for example, it increases the heart-beat, raises the blood pressure, and increases breathing rate. The cranio-sacral branch, on the other hand, acts as a check and balance to the sympathetic. It inhibits rapid heartbeat, speeds up salivary and gastric secretion (thus aiding digestion), and maintains muscular tonus. During pleasant emotional states—joy, contentment, etc.—the cranio-sacral branch of the autonomic is ascendant over the sympathetic.

In order to discover whether an intact sympathetic is necessary for emotional experiences, Cannon removed the entire sympathetic branch of the autonomic nervous system from a number of cats. This operation cut off from the viscera all sensation which might conceivably be present in fear or rage. After the operation, however, the animals all showed marked signs of anger in the presence of a barking dog: hissing, growling, retraction of the ears, and baring of the teeth.

This experiment, like that of Sherrington, indicates clearly that body sensations are not *necessary* to emotional states, even though they may be normally present. Other experiments which point in the same direction are those in which powerful drugs have been employed to bring about rapid heartbeat, trembling, blushing of the skin, rapid breathing rate—i.e., all of the changes which ordinarily are found in emotional reactions. Logically it would seem that if an individual's emotion is simply his awareness of such stirred-up states in his own body, he should feel rage or fear when so disturbed by drugs. But this result does not occur. Cannon and his associates showed in a long series of tests that adrenin (the secretion of the adrenal glands, which lie just above the kidneys) has almost the same effect upon the viscera as violently unpleasant emotions such as intense excitement, fear, and anger. When adrenin is injected into the blood stream of a normal person, the result is a decidedly stirred-up physiological state. But the subject reports no real emotion, rather he states that he feels as though he *might* be afraid without any actual fear. In connection with this finding Cannon pointed out how *alike* are the bodily or physiological changes in different emotional states, as well as in emotional and non-emotional states such as those resulting from strong muscular exercise, exposure to cold, and the like; and how *unlike* are the conscious states involved. Such observations as these seem to render the James-Lange theory untenable—at least without some modi-

fication. One case of pathological loss of sensation in a human being may be cited before leaving this topic. Dana reports the case of a woman patient, who owing to a fracture of the neck and the resulting spinal cord injury, had suffered complete paralysis of the muscles of the trunk, arms, and legs, and complete loss of sensitivity. This person lived for a year showing little or no change in personality: grief, joy, and affection seemed unaltered. Such a case can hardly be squared with the James-Lange theory if strictly interpreted.

Cannon proposed as a substitute for the James-Lange theory the hypothesis that emotions result from the interactivity of the cerebral cortex and the interbrain (thalamic regions). The interbrain, sometimes called the diencephalon, is a large mass of grey matter lying just below the cerebral cortex. It is a great sensory receptive center for impulses from the body and is a kind of vestibule for the cortex. Cannon found that when the cortex was entirely removed from a cat, the decorticated animal still showed great emotional excitement ("a sort of sham rage"), which strongly supports the view that rage, and probably other emotions as well, depend upon impulses from the subcortical regions. Following this lead Bard has shown, after some fifty operations on cats in which the cortex and various parts of the brain stem were removed at the same time, that the center for the emotional display of rage is located in the lower thalamus (i.e., part of the interbrain). According to Cannon's thalamic theory, then, at the same time that the hypothalamus (lower part) discharges upward into the cortex it also releases motor impulses which produce complex internal changes of a reflex character. In turn sensory reports of these changes in the body are sent back by way of the thalamus and so reinforce or intensify the emotional consciousness. This means that conscious emotional experience and emotional behavior are interdependent, both being aroused by nervous discharges from the hypothalamus. Cannon has cited considerable clinical evidence to show that the hypothalamus is closely connected with emotional or affective experiences. Normally this old and more primitive part of the brain is under inhibition from the cortex; and when it is "released" through disease or injury, there is often an extensive and uncontrolled emotional display (e.g., laughing or crying). The difference between one emotion and another is explained as being due to central (cortical) factors, that is, to the *meaning* which the emotion-arousing stimulus has for the individual. In the one case the stimulus may release fear and the impulse to flee; in another anger and the impulse to strike.

Cannon's thalamic theory has not been free of criticism. There is no clear experimental evidence that the hypothalamus is the chief center for the emotions. Furthermore, emotions aroused by direct stimulation of

the hypothalamus are weaker and more restricted than are natural emotions. In general, the main effect of Cannon's theory has been to switch the emphasis in studies of emotion from body changes to brain changes.

4. THE CORRELATION OF PHYSIOLOGICAL CHANGES AND EMOTIONAL STATES

Two experiments planned to find out whether definite emotional "patterns" exist physiologically and can be correlated with reported emotional states will be described in this Section. These studies, the one by Blatz and the other by Landis, are illustrative of the best experimental techniques in this field.

Blatz's experiment was concerned with discovering whether there is a definite physiological disturbance corresponding to what we call *fear*. His subjects, about twenty in all, included both men and women. They were told that the purpose of the experiment was the study of heart-rate differences over a fairly long period of time. Each subject was carefully tied in a chair and required to sit alone and blindfolded in a room for about fifteen minutes. During this period cardiac, respiratory, and electrical changes (of the skin) were carefully recorded. The force and regularity of the heartbeat were measured by an electro-cardiograph (a delicate galvanometer) and were recorded on a photographic film. Breathing records, taken by means of an electric pneumograph, were photographed on the same film. Changes in the electrical behavior of the skin, a phenomenon known as the psychogalvanic reflex,² or PGR, were obtained from the galvanometric readings. (See Fig. 28). After three sittings, Blatz found that his subjects gave consistently normal records, i.e., the slight nervousness and timidity which appeared at first due to the strangeness of the situation had been entirely lost. So true was this, in fact, that many of the subjects almost fell asleep before the fifteen-minute period was up. These preliminary records constitute what may be called a *control series*.

² The psychogalvanic reflex requires a word of explanation. If electrodes are attached to two different points on the skin and are connected with a sensitive galvanometer, an electric current (as shown by deflections of the galvanometric needle) will be found to flow between the two points. This current indicates the difference in electrical potential between the two points, and is found to increase markedly during emotional states; hence it is often taken as an index of the intensity of the emotional state. The galvanic deflection is called the psychogalvanic reflex. It is due presumably to changes in sweat-gland secretion. The sweat-glands are controlled by the sympathetic branch of the autonomic nervous system, which, as we have seen (p. 151), is active in intense emotional reactions. The assumption, then, is that fear or anger, say, cause an increased outpouring of sweat secretion; this sweat, in turn, lowers the skin resistance to the electric current, thus giving an increased galvanic deflection.

The chair used in this experiment was hinged in front so that when suddenly released it would fall over backward, carrying the subject with it. A powerful door-check stopped the fall gradually after the chair had passed through an arc of sixty degrees. At the fourth sitting, Blatz allowed his subject without warning to fall over backward. This backward drop was intended to set up a fear response, so that records taken at the time would give a physiological picture of that emotion which we subjectively call fear. That the fall really caused fear can hardly be doubted. All of the subjects reported genuine fear at first, followed later on, oftentimes, by anger or amusement. In addition, they struggled, yelled out, called the experimenter by name, and in various ways gave evidence of emotional upset. The effects produced by the fall may be briefly summarized as follows (records were taken before, at the time of, and six minutes after the fall):

1. Breathing was much disturbed. As the chair fell over backward the subjects tended to "catch their breath," and then exhale in short gasps.

The fraction $\frac{\text{Inspiration time}}{\text{Expiration time}}$, called the "respiratory ratio," was immediately *increased*. That is to say, the breathing-in became relatively much longer than the breathing-out, the ratio I/E increasing from a normal of about .70 to as high as 4.00.

2. Heart rhythm was irregular. The heart-rate was markedly accelerated at first by as much as twenty beats (from 84 to 104 beats on the average). This initial increase was usually followed by a retardation and then another period of acceleration.

3. In the psychogalvanic reflex there were wide deflections. These began from one-half second to three seconds after the drop and ranged from one to ten units over the scale.

In addition to these organic disturbances, there were also gross muscular movements, struggles to escape, reflex movements, and thrashings of the arms and legs, and yelling. In later tests, when the subject expected a fall (having been forewarned), the physiological effects were less intense than when the fall was unexpected. Adaptation was rapid, the second and third falls producing much less upset than the first.

This experiment, to be sure, reveals a definite picture of physiological disturbance but it offers no comfort to the James-Lange theory. There is no pattern or regularity in the response which is constant from person to person or for a given individual from time to time. Furthermore, there is too much variability in all of the measures taken to indicate the presence of a clear-cut physiological state which can confidently be labeled "fear" or "terror."

In his experiment, Landis investigated the effects of severe emotional

upset upon respiration, blood pressure, gastro-intestinal activity, and basal metabolism. Respiration, both thoracic (chest) and abdominal, was measured by pneumographs. Blood pressure, which can best be described briefly, perhaps, as the pressure in the arteries resulting from the force of the heartbeat and from arterial tension was obtained by means of a sphygmograph³ attached to the upper arm. Gastric contractions were obtained by having the subject swallow a small balloon attached to a rubber tube, so that changes in the size of the balloon resulting from stomach contractions were recorded directly. Rectal contractions were secured by using a similar balloon arrangement placed in the rectum. Basal metabolism (roughly, the amount of carbon dioxide given off in breathing) was obtained by having the subject breathe into a specially designed apparatus. All of these measures of physiological change were recorded simultaneously by means of tambours and writing points upon a moving smoked paper.

Landis used three subjects in this experiment. After normal records had been taken over a period of three weeks, the following schedule was begun. Each subject went without food for forty-eight hours and without sleep for thirty-six hours. At the end of this time an electric shock was administered which was "as much as the subject could bear without struggling," and which lasted until the subject signaled that he could endure the pain no longer. Each subject was then given a stimulant, and records were taken hourly for a five-hour period and daily afterward for five days. Landis's results in brief were as follows:

1. *Basal Metabolism.* Basal rate increased about 30 per cent on the average (with much variability) during the time of anticipation before the shock; it then fell off rapidly until, after six to eight hours, it was practically normal again. Anger, which occurred occasionally, was accompanied sometimes by an increase and sometimes by a decrease in basal metabolic rate.

2. *Blood Pressure.* There was a rapid rise in blood pressure upon the onset of the shock, together with a rapid and irregular heartbeat. Upon continuation of the stimulus the heartbeat became more regular and the pressure decreased. Landis states that these changes are quite similar to the shock symptoms following a surgical operation.

3. *Respiration.* Breathing was faster and more shallow, with some tendency toward gasping.

4. *Gastric and Rectal Contractions.* The electric shock stimulus had

³ This instrument consists essentially of a rubber sleeve attached to the arm. When it is inflated to the point where the flow of blood is cut off, the blood pressure is measured either graphically or on a scale, by a tube leading off from the inflated sleeve.

a variable effect on gastric contractions. In one subject gastric contractions were marked, while in another they ceased temporarily. The third subject could not retain the apparatus, and hence gave no record. Rectal contractions ceased in all subjects upon the onset of stimulation.

On the day following the experiment, each subject wrote out an account of his experiences, attitudes, and feelings. All stressed fatigue, nausea, occasional fits of anger, irritability and lack of coördination. There was little or no awareness of specific physiological changes, and no agreement as to the specific emotion felt. As in Blatz's experiment, no evidence appeared to indicate that a regular pattern of physiological change invariably accompanies a reported emotional state.

The only behavioral pattern which seemed to correspond *regularly* to a given subjective state was that of *surprise*. Surprise is the immediate response to any sudden and intense stimulus; it has been better named, perhaps, the "startle pattern." When studied with the aid of high-speed photography the startle pattern has been found to comprise two main elements: the eye blink and a complex, but identifiable, facial-bodily reaction. These two responses are reflexive and involuntary, and constitute a definite pattern of stereotyped behavior. "Startle" is the only emotional state which fits into the James-Lange definition of an emotion as an *invariable* complex of bodily changes.

5. JUDGING EMOTIONS FROM FACIAL EXPRESSIONS

Even if reported emotional states are not highly correlated with specific physiological changes in the body perhaps emotions are related to facial expression or appearance. One of the earliest attempts to discover how accurately one can judge the probable emotional state of a person from a photograph is that of Feleky. To each of one hundred judges this investigator submitted eighty-six photographs of herself, posed to represent a wide range of emotional states, together with a list of common names for emotions. The judges were instructed to write on a sheet opposite the number of each photograph the emotion which they believed it portrayed. The emotions which were represented ranged from fear, rage, and suspicion to surprise, sympathy, and religious feeling. Perhaps the most striking first impression which one gets from Feleky's data is the wide range of emotions attributed to each pose. Closer examination shows, however, that—allowing for synonyms and variations in descriptive names—there is rather remarkable agreement on many of the poses. The following table gives the percentage of times a selected group of posed expressions was judged correctly or nearly correctly:

<i>Pose</i>	<i>Called by Judge</i>	<i>Percentage of Judgments</i>
Happiness	Happiness, mirth	90
Surprise	Surprise, happiness	82
Fear	Fear, suffering, anger	91
Anger	Anger, fear, suffering	73
Disgust	Disgust, contempt	75
Contempt	Contempt, disgust	93

This experiment is limited, of course, in its practical application, since one never strikes and holds a pose when normally upset. But it does indicate that we often infer an individual's emotional state quite accurately from a photograph.

Landis has performed an experiment designed to discover whether reported emotions are accompanied by definite and easily recognized facial expressions. A series of photographs were taken of his subjects while they were undergoing various emotion-producing situations. There were seventeen of these situations in all, several of the more striking being: smelling a bottle containing ammonia, falsely marked lemon, looking at pornographic pictures; decapitating a live rat; severe electrical shock. Twelve women, twelve men, and one boy acted as subjects. In order the better to analyze the kind and amount of movement exhibited by the facial muscle groups, these were traced out upon the face of each subject with burnt cork to make them show up better in the photographs.

The subjects characterized the emotional value of the situations in various ways, by profanity, outcries, and verbal descriptions. Disgust, anger, surprise, and sex excitement were reported often enough to be studied in connection with the photographs of facial expressions taken during these states. After many comparisons Landis writes: "With no verbal report of a given emotion did a muscle, group of muscles, or expression occur with sufficient frequency to be considered characteristic of that emotion. There is no expression typically associated with any verbal report."

Perhaps this negative finding is to be explained in part by the fact that (1) emotional states changed rapidly during the experiment and were often not very genuine—smiling was the most common reaction reported; (2) the photographs were small; and (3) the markings on the face were probably confusing rather than helpful. At any rate, it appears that our inferences of emotional states from photographs are based upon the expression conveyed by the whole face rather than by parts of the face or specific muscle groups.

A study of children's judgments of posed facial expressions shows that youngsters judge pleasant and fairly simple emotions better than they do unpleasant and more sophisticated states. The pose for laughter was

judged correctly 70 per cent of the time at ages 3 to 4 years, and 100 per cent accurately by age 10. The age-levels at which certain emotions were correctly judged by at least 50 per cent of the children examined are as follows:

<i>Pose</i>	<i>Age</i>
Laughter	below 3 years
Pain	5-6 years
Anger	7 years
Fear-horror	9-10 years
Surprise	11 years
Contempt	above 14 years

Contempt, which was hardest for children, was easy for adults as shown by Feleky's data (p. 158). Facial expressions typical of many emotional states are fairly standard in our society and are, in fact, indices of social communication. Failure to name correctly the more "adult" emotions such as contempt, suspicion, jealousy and the like may mean simply that children have no *names* for these emotions. Hence they are classified simply as unpleasant.

6. THE PSYCHO GALVANIC REFLEX

Another problem which arises in studying the emotions is that of the extent to which the intensities of an individual's various emotional states are correlated. Does the person who gives an intense fear response to a fear-provoking situation, for instance, tend to react in the same highly emotional way to situations provoking anger, joy, or grief? Several investigators have attempted to answer this question, using the psychogalvanic reflex as an indicator of emotional intensity. The psychogalvanic reflex has been described on p. 154 as an electrical phenomenon which results from a difference in potential between two points on the skin. The electrical response is dependent upon changes in sweat-gland secretion which are, in turn, a result of activity in the sympathetic branch of the autonomic nervous system (p. 152). Since the autonomic nervous system is known to be active in emotional behavior, the psychogalvanic reflex or PGR thus becomes, potentially at least, an index of emotional states. The sketch of a simple setup for securing the galvanic reflex is shown in Figure 28, and in Figure 29 is pictured an ordinary D'Arsonval galvanometer.

In one experiment with the PGR eleven different stimuli were employed, such as pricking the subject with a needle, sounding an automobile "klaxon" horn suddenly, and ringing a bell. All of these, possibly excepting two or three, were calculated to surprise or annoy the subject by arousing anger, pain, or mild fear. The subject's response to each situation was

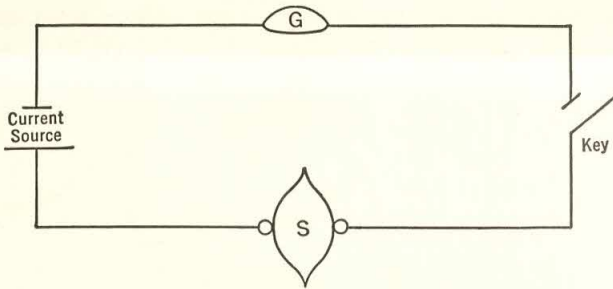


FIG. 28

SIMPLE SKETCH TO SHOW HOW THE PSYCHOGALVANIC REFLEX (PGR) MAY
BE MEASURED

G is the galvanometer and S is the subject, who is in circuit with G and with a source of current, usually of two to four volts. When the circuit is closed by the key, there will be an initial galvanometric deflection on the scale (see Fig. 29), which is recorded by the experimenter. When the subject is stimulated, the deflection measured from this point is an indicator of the PGR. A Wheatstone bridge is often employed to bring the initial deflection back to zero, so that all PGR's may be taken from this reference point.

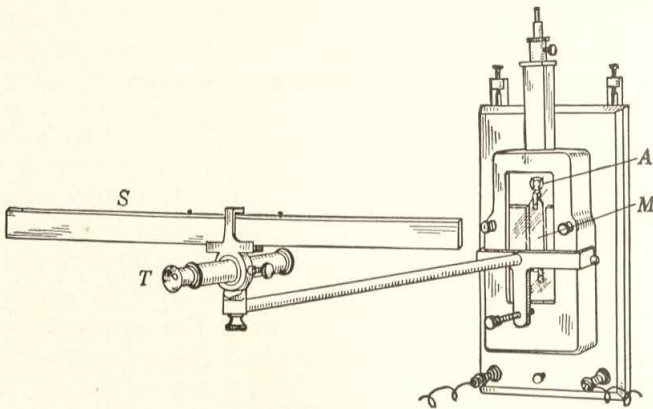


FIG. 29

D'ARSONVAL GALVANOMETER

Changes in the subject's resistance produced by emotional stimuli cause movements of the magnet, M, to which is attached a small mirror, A. As A turns to the right or left, these deflections may be read from the scale, S, by the experimenter, through the telescope T.

measured by the galvanometric reading. The galvanic response to the sound of a horn is shown graphically in Figure 30. In accordance with the extent of the psychogalvanic response, the thirty subjects were ranked in order of merit for the eleven situations. The correlation or correspondence of these rankings ranged from 0 or negative (just no relationship) to .80 or more (close relationship), with an average at .58.⁴ In the case of

⁴ For meaning of correlation, see p. 269.

the high correlation of .80, this very close correspondence was almost certainly due to the identity of the two stimuli involved; the first was the sudden sound of an automobile horn, and the second the same stimulus repeated.

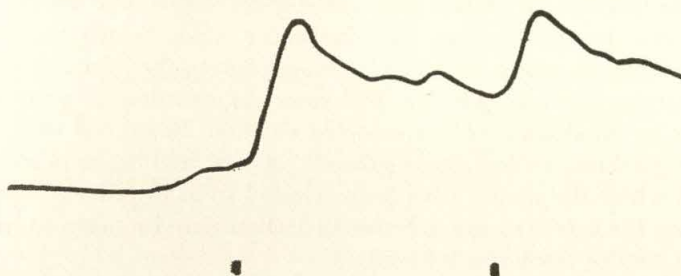


FIG. 30

SHOWING THE GALVANIC RESPONSE (PGR) TO THE SOUND OF A HORN

The dots indicate the sound stimuli. (From Wechsler, *The measurement of emotional reactions: researches on the psychogalvanic reflex*, Arch. Psychol. 1925, p. 26.)

This experiment is certainly not conclusive, but within the limits of the laboratory situation offers little evidence that emotions are highly correlated—that persons tend to be *generally* emotional or unemotional. Other experiments have also found the same tendency toward specificity of emotional reactions. And these experimental findings are substantiated to a considerable degree by everyday experience. A few people may be fairly described as “choleric”—easily disturbed or excited by a variety of stimuli, some relatively mild. Others, more “phlegmatic,” react mildly even to intense and demanding stimuli. But these are exceptions. Most people fall in between these two groups and the “generally” excited and “generally” placid constitute small groups which fall at the extremes of an emotional scale. Only when situations are closely alike or the emotions elicited quite similar are intercorrelations high. In a number of situations, for example, *all* designed to provoke surprise or startle, Patterson found agreement between the *intensity* of the PGR in various situations to be expressed by correlations of from .53 to .88.

The question of what conscious state or states most regularly accompany the PGR has been raised by Landis and Hunt. These authors refer to the PGR as the GSR (galvanic skin response), a better term, perhaps, because it describes rather than interprets the actual phenomenon. Landis and Hunt presented to their subjects various stimuli intended to arouse amusement, fear, sex emotion, unpleasantness, and other emotional states. Each observer was asked to report his subjective feelings or mental state after each response. When reports were tabulated, the PGR was found to

be most nearly related to *tensions*—fear, anger, startle, and the like. But the galvanic response occurred under such a wide variety of conditions that it was clearly not an index of any one conscious state.

The main objection to the use of the galvanometer as an index of emotion is the fact that the PGR is not invariably correlated with activity in the sympathetic nervous system. Skin resistance, upon which the psychogalvanic reflex depends, is known to change during the course of the day, while fatigue, deep respiration, and muscular exertion all provoke deflections in the absence of any reported emotion. Emotional states, however, do produce marked psychogalvanic reflexes: and hence in an experiment in which the stimuli have been selected so as to provoke emotional reactions, the extent of the galvanic deflection may be taken as an indicator of *relative* emotional intensity.

Cannon conceived of the sympathetic branch of the nervous system as performing an *emergency* function. When animal or man is confronted by sudden danger, or is trapped and must attack or escape quickly, the sympathetic system “takes over,” according to this view, and the organism is “stripped for action.” Emotion, then, becomes a “reinforcer” of bodily action—a means of supplying additional energy to the body muscles during the emergency. The experimental connection established between sympathetic activity (via the PGR) and reported conscious emotional states of tension, startle, confusion, and the like fits in well with Cannon’s “emergency theory” of the emotions.

7. THE ASSOCIATION METHOD IN EMOTIONAL DIFFICULTIES

The experiments presented above show some of the possibilities of the method of expression. A good illustration of the method of impression is furnished by the free-association test. Used as a means of diagnosing emotional difficulties or complexes, the association test method was first employed by the Swiss psychiatrist Carl Jung. Jung’s method was to present to his subject or patient a set of words (usually 100), to each of which the patient was instructed to give as quickly as possible the first word or association which occurred to him. The stimulus words were selected to cover a wide range of situations, including many which would presumably have emotional value for the subject. These so-called “critical words” were mixed in with many different or innocuous words. By way of illustration Table XX gives a selection of twenty-five critical words from Jung’s list. The theory underlying the association method is that extreme timidity, embarrassment, useless fears, anxieties, worries, and the like which occur in nervous or neurotic persons, and to a lesser degree in normal persons, center around forgotten and little-understood emotional epi-

TABLE XX

A SELECTION OF TWENTY-FIVE CRITICAL WORDS FROM JUNG'S LIST
OF 100 ASSOCIATION WORDS

dead	old
to dance	to beat
sick	to wash
angry	to fear
to swim	brother
pity	false
to die	anxiety
to pray	to kiss
money	bride
despise	pure
unjust	contented
to marry	woman
ridicule	

sodes in the person's life. Those words in the list which revive these occurrences, or "tap" complexes, should therefore provoke highly personal or emotional associations accompanied, say, by laughter or blushing. Lengthened reaction time, repetition of the stimulus word, silly or far-fetched answers, or an entire lack of response are interpreted as an attempt to avoid unpleasant associations connected with the stimulus word. Such reactions are called "complex indicators." A complex, as the psychoanalyst uses the term, denotes a group or constellation of ideas centered around some particular episode or theme which possesses emotional significance to the subject. An individual much excited or disturbed over religious matters is said to have a "religious complex"; other persons are described as having sex complexes, or inferiority complexes, the latter term being used to cover feelings of inadequacy of various kinds.

On the principle of psychic determinism, every reaction is explicable in terms of its antecedents, since all of these are inevitably linked up with the response. So it is argued that any word tied up in the subject's experience with an original emotional upset will, if followed up through free association, finally lead back to the source of his difficulty. To illustrate, a highly emotional or lengthened response to the word *ridicule* might enable the psychologist to discover by further questioning that the examinee's "touchiness" regarding ridicule grew out of his having a peculiarly shaped nose or a speech impediment. This defect had led others to tease and torment him so that feelings of inadequacy and the misery arising therefrom had interfered with his normal life. In Table XX the words *kiss*, *bride*, *pray*, *ridicule*, *anxiety*, for instance, might conceivably lead to reactions which would be considered complex indicators.

The validity of a complex indicator as a revealer of emotional stress has been studied in several ways. Checking one indicator against another is perhaps the best procedure; and this method has revealed considerable agreement among certain responses. Repetition of the stimulus word and misunderstanding of the stimulus word are among the best indicators; long reaction time and irrelevant responses are also good but somewhat less valuable than the first two. Two indicators occurring together are more "diagnostic" than a single one. The emotionally toned or critical words differ from one investigation to another for different subjects, but they are usually concerned with love, anger, injustice, ridicule, pity, danger, and death.

The association method in one form or another has been much used by clinical psychologists. Perhaps the chief objection to association tests as indicators of troublesome and unconventional thoughts is that they may easily prove too much. Nearly every person has frequently been worried about religious matters or sex adjustments, or has felt inferior, and if sufficiently prodded, can probably reveal several "complexes." This is particularly true of the sex life, since present-day society curbs and limits freedom of expression in such matters.

Extensive use of free association has been made by Kent and Rosanoff in their comparison of the responses given by the normal and the insane. These investigators drew up a list of 100 words, some having probable emotional value (i.e., covering the usual worries and anxieties), and some being presumably indifferent such as *table*, *chair*, *stove*. This list was administered to 1,000 normal and 247 insane adults, the responses to each word being carefully tabulated. Probably the most striking result obtained from this mass of material was the large percentage of "individual" responses given by the insane as compared with the normal. If a person gives a response not duplicated by any one else in the group, it is considered to be an "individual response." The normal group gave about 7 per cent individual responses, the insane group 27 per cent. A number of individual responses is considered by the authors to be indicative of eccentric thinking or other peculiarity; while a number of "common" responses (those given by others) indicates normality in the sense of conformity to the standards of the group. Each response, together with the frequency of its occurrence, has been tabulated for each of the 100 words in the list. From these tables one can determine how commonplace or how exceptional his associations are.

Frequency tables have also been prepared giving the responses of 1,000 school children, ages 9-12, to 100 stimulus words—90 of these taken from the Kent-Rosanoff list. In general, children's responses are more *concrete* than those of adults. They tend to be in terms of use, or to tell something about the idea expressed by the stimulus word. To the stimulus word

"table," for example, 36 per cent of the children answer "eat," while only 5 per cent of adults say "eat," 30 per cent answering "chair." To the stimulus word "man," children reply "work," adults, "woman"; to "soft," the children say "pillow," the adults "hard." *Coördinate* responses (man—woman, table—chair) and *opposites* (dark—light, deep—shallow) are given more often by adults than by children. These differences probably reflect the adult's larger vocabulary, and the tendency for adults to think in abstract terms to a greater degree than children.

The word association test has proved useful in mental hospitals as a research tool and as an aid in diagnosis. In their studies of clinical tests, Rapaport and his associates have employed a set of 60 stimulus words—called the Orbison List after its author. This list includes words relating to home and family as well as words having unpleasantly aggressive and sexual connotations. Twenty of the words are classified as "traumatic," that is, as being likely to arouse emotional disturbances. The test is administered first as a free-association test and later the list is read to the patient who is asked to reproduce his original answers.

Responses are studied and classified in various ways and are examined for indicators of emotional disturbance (p. 167). Figure 31 shows the percentages in each of three groups who revealed 15 or more associative disturbances. These groups were (1) schizophrenics and depressives, (2) neurotics, and (3) normals. Increasing maladjustment is evident as we go

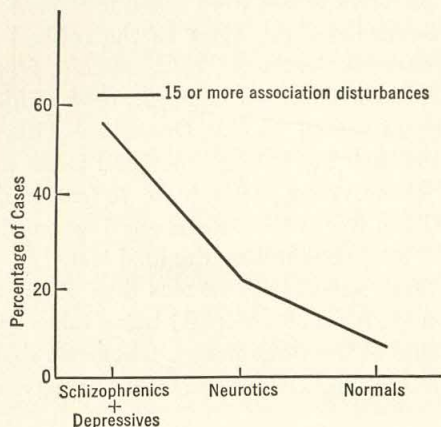


FIG. 31
ASSOCIATION DISTURBANCES IN THREE GROUPS
(From Rapaport, 1946. The Year Book Publishers, Inc.)

from the normals to the psychotics. Nearly 60 per cent of the psychotics gave 15 or more disturbed responses as compared with not quite 10 per cent of the normals. The neurotics fall in between the two being nearer to the normals than to the psychotics.

8. ATTEMPTS TO DISCOVER GUILT OR DECEPTION

From the experimental point of view, the "detective" use of the free association test is one of the most direct applications of the *method of impression*. As its name implies, this method is intended to discover the "real" culprit from among several persons suspected of a "crime." In the psychological laboratory, the "crime" is usually a stunt of some kind through which one person out of several possible candidates is put. It is so arranged that neither the experimenter nor the class knows who is the "guilty" person. The object of the association test is to select this guilty individual from among the several suspects by means of his telltale associations. Usually these responses will be abnormally long or emotionally tinged (accompanied by laughter or embarrassment), or they will bear directly upon the stunt, thus giving the culprit away.

Jung has employed the word-reaction method in the following experiment, which may be used as an illustration. The supervisor of a hospital reported to him the theft of a pocketbook from one of the nurses in her charge. The purse contained a fifty-franc note, one twenty-franc piece, some centimes, a small silver watch-chain, a stencil, and a receipt from Dosenbach's Shoe Shop in Zurich. The purse had been taken from a clothes-closet in which it had been placed by the nurse. Owing to various circumstances which need not be described in detail, suspicion narrowed down to three nurses, all of whom were asked to submit to the association test. The critical words were the name of the robbed nurse, *cupboard, door, open, key, yesterday, banknote, gold, seventy, fifty, twenty, money, watch, pocketbook, chain, silver, to hide, fur, dark reddish leather* (color of the purse), *centimes, stencil, receipt, Dosenbach*. Other words not bearing directly upon the theft but having emotional value were *theft, to take, to steal, suspicion, blame, court, police, to lie, to fear, to discover, to arrest, innocent*. These critical words were distributed among twice as many indifferent words, the total constituting the final test.

To each of the three nurses the test was then given, the response and the time of reaction (in fifths of a second) being taken for each word. The median reaction times of the three nurses, whom we shall designate A, B, and C, to the indifferent and the critical words are given in the following table:

REACTION TIME (FIFTHS OF A SECOND) OF A, B, AND C TO THE
INDIFFERENT AND CRITICAL WORDS

	A	B	C
Indifferent words	10	11	12
Critical words	16	13	15
Difference	<u>6</u>	<u>2</u>	<u>3</u>

Although A's "normal" reaction time—as shown by her responses to the indifferent words—is the shortest of all, she is considerably slower than either B or C in replying to the critical words. This, of course, is mild though surely not conclusive evidence against A. Jung next computed the number of "imperfect reproductions" given by each nurse. An imperfect reproduction or reaction is one which is haltingly or stumblingly given, with repetition or evident emotional upset. Such responses may grow out of an association of strong feeling-tone aroused by the critical word which is carried over to several succeeding responses. The subject, so to speak, becomes "rattled," and gets more flustered as the experiment goes on. The result of this tabulation showed that A gave 65 per cent imperfect reactions, B 56 per cent, and C 30 per cent. The actual responses were distributed as follows:

NUMBER OF IMPERFECT REACTIONS GIVEN TO INDIFFERENT AND
CRITICAL WORDS BY A, B, AND C

	A	B	C
Indifferent words	10	12	11
Critical words	19	9	12
Difference	<u>9</u>	<u>3</u>	<u>1</u>

A has an excess of 9 responses to the critical words, B has 3, and C has only 1. By this test, then, suspicion again points to A. Still another check was made in terms of the percentage excess for each subject of "complex indicators" given to the critical words, over and above those given to the presumably indifferent stimuli. A's excess is 100 per cent, B's 0, and C's 50 per cent. On the basis of these statistical results and upon careful study of the character of the responses, Jung decided that the greatest suspicion fell on A. Confession by A later on confirmed this judgment.

The *method of expression* has also been used in attempting to discover guilt or deception. Both Marston and Larson have employed changes in blood pressure, as evidence of lying or deception, with much success. Marston believes that a rise in blood pressure of from 8 to 12 millimeters in conjunction with a response to a possibly compromising question is sufficient to indicate that the subject is lying. Larson, who first used the association test method and later the method of direct questioning, has employed blood-pressure rises and changes in respiration with success in several cases involving actual crimes. It is probable that blood-pressure changes are indicative of deception because the attempt to deceive causes general excitement which in turn raises the blood pressure. Blood pressure, however, is affected by many factors and its use in crime detection is still experimental to a high degree.

Perhaps it need not be pointed out that neither free association nor any of the other so-called "lie-detector" methods such as blood-pressure changes and PGR can pick out a guilty person with certainty. The habitual liar, for instance, is undisturbed in telling untruths, while a deluded or fanatical person may convince himself that he is sincere. Nor can these methods distinguish between one who is guilty and one who is not guilty but possesses guilty knowledge. They do yield valuable clues in many cases, however, and are often useful in leading to confession.

9. PRESENT STATUS OF THE STUDY OF EMOTIONS

In the first part of the chapter we outlined four views—representing somewhat different emphases—of what constitutes an emotion. According to the *first* view, an emotion is a mental or conscious phenomenon largely independent of the body changes which are its usual but somewhat incidental accompaniments. In the *second* view, an emotion is considered to be the individual's perception of the sensory changes arising from glandular, muscular, and circulatory reactions in his own body (this is the James-Lange theory). The *third* view regards an emotion as a felt impulse or set toward a certain type of activity plus a mass of sensations resulting from disturbed body states. The *fourth* view is the behavioristic dictum that an emotion is simply the reaction pattern itself, the conscious state, if admitted at all, being of no consequence.

Let us now attempt to evaluate these views as best we may in the light of the experimental evidence at our disposal. The *first* view, that an emotion is simply a conscious or mental state, is clearly too narrow. The part played by body changes in preparing the individual to meet danger or some other emotion-producing situation is too real to be ignored. Furthermore, we cannot omit from a treatment of emotional reactions the rôle of the skeletal muscles nor the enormous importance of the glands in emotional states. The *second* view—that expressed by the James-Lange theory—is largely untenable in view of the experiments of Sherrington and Cannon, though to this theory should go the credit for being the first to recognize the reflexive component in emotion. The *fourth* view (the extreme behavioristic) seems to most psychologists to involve a distortion and even a suppression of the truth. Certainly experiments carried out with the best techniques now available have failed to find for a given emotion either a definite physiological pattern or a very characteristic behavior response.

This leaves us, finally, the *third* viewpoint. This, on the whole, is the most satisfactory, since it provides both for the conscious side of the emotion and for the body changes which accompany it. The "pattern"

for an emotion, according to this view, does not lie in the mass of physiological changes, which vary little from one emotion to another, but rather in the *impulse* or *set* of the organism. In fear, for example, the set is to get away—to escape the influence of the stimulus—while in joy the impulse is to continue the influence of the stimulus. Usually the impulse precedes the bodily changes, which are often biologically valuable as preparatory reactions. This view is consistent with the emergency theory of the emotions described on p. 153.

This chapter has attempted to show that as a result of experimental attack there is a respectable body of knowledge regarding the emotions. Much is now known about the nervous and physical basis of emotional states; many of the complex physiological, circulatory, and electrical changes accompanying emotions have been measured; and techniques have been developed for investigating the causes of certain abnormal emotional states. A great deal remains to be done. Future research will further define and differentiate emotions both on the physiological and psychological side. And genetic studies will aid in showing how emotions develop, how they become organized, and, possibly, how they may be better controlled for the benefit of the individual and of society.

BIBLIOGRAPHY

1. Bard, P. The neuro-humoral basis of emotional reactions. Chapter 6 in Murchison, C. (Ed.). A handbook of general experimental psychology. Worcester, Massachusetts: Clark University Press, 1934 (esp. pp. 290–296).
2. Blatz, W. E. The cardiac, respiratory, and electrical phenomena involved in the emotion of fear. *J. exp. Psychol.*, 1925, 8, 109–132.
3. Cannon, W. B. The James-Lange theory of emotions: a critical examination and an alternative theory. *Amer. J. Psychol.*, 1927, 39, 106–124.
4. Cannon, W. B. Bodily changes in pain, hunger, fear and rage. New York: D. Appleton and Company, 1929.
5. Chappell, M. N. Blood pressure changes in deception. *Arch. Psychol.*, New York, 1929, 17, No. 105.
6. Crosland, H. R. The psychological methods of word association and reaction time as tests of deception. University of Oregon Publications, Psychol. Series, 1929, 1, No. 1.
7. Dana, C. S. The autonomic seat of the emotions: a discussion of the James-Lange theory. *Arch. Neurol. Psychiat.*, Chicago, 1921, 6, 634–639.
8. Feleky, A. M. The expression of the emotions. *Psychol. Rev.*, 1914, 21, 33–41.
9. Gates, G. S. An experimental study of the growth of social perception. *J. educ. Psychol.*, 1923, 14, 449–461.
10. Hull, C. L., & Lugoff, L. S. Complex signs in diagnostic free association. *J. exp. Psychol.*, 1921, 4, 111–136.
11. Jung, C. G. The association method. *Amer. J. Psychol.*, 1910, 21, 219–269.

12. Jung, C. G. *Studies in word-association*. New York: Moffat, Yard & Company, Inc., 1919, chapters 2, 5, 7, 10.
13. Kent, G. H., & Rosanoff, A. J. *A study of association in insanity*. *Amer. J. Insanity*, 1910-11, 67, 37-96, 317-390.
14. Landis, C. *Studies of emotional reactions*. II. General behavior and facial expression. *J. comp. Psychol.*, 1924, 4, 447-509.
15. Landis, C. *Psychology and the psychogalvanic reflex*. *Psychol. Rev.*, 1930, 37, 381-398.
16. Landis, C. *Studies of emotional reactions*. V. Severe emotional upset. *J. comp. Psychol.*, 1926, 6, 221-242.
17. Landis, C. *Emotion*. II. The expression of emotion. Chapter 7 in Murchison, C. (Ed.), *A handbook of general experimental psychology*. Worcester, Massachusetts: Clark University Press, 1934.
18. Landis, C., & Hunt, W. A. The conscious correlates of the galvanic skin response. *J. exp. Psychol.*, 1935, 18, 505-529.
19. Landis, C., & Hunt, W. A. *The startle pattern*. New York: Farrar & Rinehart, 1939.
20. Larson, J. A. The cardio-pneumo-psychogram in deception. *J. exp. Psychol.*, 1923, 6, 420-454.
21. Lashley, K. S. The thalamus and emotion. *Psychol. Rev.*, 1938, 45, 42-61.
22. Marston, W. M. Systolic blood pressure symptoms of deception. *J. exp. Psychol.*, 1917, 2, 117-163.
23. Patterson, E. A qualitative and quantitative study of the emotion of surprise. *Psychol. Monog.*, 1930, 40, No. 181.
24. Rapaport, D. *Diagnostic psychological testing*, Vol. II, 1946, Menninger Clinic Monog., Series No. 4, chapter 2.
25. Sherrington, C. S. *The integrative action of the nervous system*. New Haven, Connecticut: Yale University Press, 1906, 259-262.
26. Wechsler, D. The measurement of emotional reactions: researches on the psychogalvanic reflex. *Arch. Psychol.*, New York, 1925, 12, No. 76.
27. Woodrow, H., & Lowell, F. Children's association frequency tables. *Psychol. Monog.*, 1916, 22, No. 97.
28. Woodworth, R. S., & Marquis, D. G. *Psychology*, 5th ed. New York: Henry Holt and Company, 1947, 417-439.

EXPERIMENTAL APPROACHES TO PERSONALITY STUDY

1. THE MEANING OF PERSONALITY

THE TERM *personality* is one of those omnibus expressions difficult to pin down and define precisely. The merchant wants salespeople who are healthy and well trained and who, at the same time, possess "good personality." In a salesperson, a good personality includes behavior activities ranging all the way from character traits such as honesty and sincerity to pleasant social behaviors such as courtesy and good manners. Not all of these attributes, to be sure, are regarded as equally important, but all are felt to be useful in some degree. Terman wrote that "Above an I.Q. level of 140, adult success is largely determined by such factors as social adjustment, emotional stability, and drive to accomplish"—that is, by personal qualities. Often we are more impressed by how a man does his job than by how much he knows about it. In fact, more often than not we attribute the success of a physician, politician, waiter or garage mechanic to patience, tactfulness and a friendly attitude rather than to intelligence and training. Perhaps we cannot identify the specific personality variables which contribute to success in a given case, but we are convinced of their importance nonetheless.

Running through the various descriptions of personal and social behavior is a common core of agreement. Thus personality is thought of as related to but not identical with intelligence; and as being closely identified with social ease and emotional balance, that is, with how one affects people and how one is affected by them. Personality, then, is *style* or *form* of behavior. What stamps an individual as a distinctive "personality" is the organization of his behaviors, habits and attitudes rather than the mere existence of certain specific traits. Just as the anatomical relationships of a man's features serve to make his face peculiarly his own, so the ways in which his personal traits interact and supplement each other render an individual's behavior unique. In short, personality is revealed

in the distinctive and characteristic ways in which a person reacts to everyday situations.

2. FREUD'S INFLUENCE UPON THE STUDY OF PERSONALITY

Sigmund Freud was born in 1856 in what is now Czechoslovakia. From the age of four he lived in Vienna, where he was trained to be a physician and where, from 1886 until 1938, he practised medicine. Freud is known almost universally as the founder of psychoanalysis, which is, at the same time, a system of mental therapy and a "school" of psychology with a carefully formulated theory of personality structure. Psychoanalytic concepts have exerted a profound influence upon modern thinking in child and in social psychology, in sociology and in anthropology.

In his practice and teaching, Freud was primarily concerned with the interplay or *dynamics* of human motives and with the development of personality. His study of motivation in persons who were mentally disturbed led him to theorize concerning the drives and emotional life of normal people, and to go on from here to many fundamental social, cultural, and even religious problems. Freud believed that our motives are usually (perhaps always) "unconscious" and are thus not a part of everyday awareness. Consciousness, in psychoanalytic theory, is analogous to the small part of an iceberg which protrudes above the surface of the sea while the bulk of ice (the "unconscious") lies far below the surface. This submerged (or unconscious) part of mental life is full of stresses and conflicts, and the goal of psychoanalysis is to bring these conflicts to the surface (into consciousness) so that they may be understood and perhaps alleviated.

In brief, Freud's theory of personality development runs as follows. The child, he says, is dominated by two powerful instincts: the *ego instinct* and the *sex instinct*. The first is identified with self-preservation and is active in hunger, fear, self-assertion. The second includes not only the urge toward reproduction and normal sexual desire but almost any kind of pleasure seeking. This view of motivation is expressed somewhat picturesquely by Freud in the following quotation:

The poet philosopher Schiller once said

"Until philosophy sublime,
Supremely rules the course of time,
The World, in oldest fashion,
By hunger moves, and passion."

Hunger and passion are two very powerful agents! . . . They want to be satisfied, that is, they endeavor to create such situations whereby the needs of our body are gratified. (Ref. 12, p. 68.)



SIGMUND FREUD

1856-1939

By "hunger," it is clear that Freud meant not only desire for food but the ego instinct—in general, the aggressive striving for mastery. Passion he takes to include all that pertains to sex in a wide sense. Together these two, hunger and sex, are thought of as shaping the course of personality development.

Freud believed that every growing child passes through three growth periods or stages. In the infant it is difficult to separate sexuality from the gratification of sensory and bodily needs, provided by the affectionate care of the parents. The boy of 4 or 5, says Freud, is apt to be jealous of his father's attentions to his mother, who is his chief source of affection and security. The girl of the same age encounters a somewhat similar problem, except that her affection is usually directed toward the father who provides security, while she is jealous of her mother. During and before puberty, Freud believed that boys and girls go through a second or "homosexual" stage in which boys are interested mainly in boys and girls in girls. This behavior is in a sense forced upon the adolescent for custom and convention tend to frown upon close social relations at this age between the sexes while encouraging friendship between persons of the same sex. By the age of 16 to 20, courtship between the sexes is permitted and expected, and the third or heterosexual stage is reached.

In Freud's view a neurosis or severe mental maladjustment arises from the blocking or diversion of the sex drive during infancy or childhood. The energy of the ego or of sex drive cannot be repressed: it can only be redirected or perhaps exhausted. The peculiarities in the behavior of the neurotic person, his anxieties, feelings of guilt, tensions, compulsions, and physical complaints result (according to psychoanalytic theory) from "unconscious" conflicts between incompatible motives or between the sex motive and the restrictions imposed by society.

While Freud's concepts of motivation and of personality development are central in present-day psychiatry, their validity for the most part depends upon clinical reports and observations rather than upon experimental verification. However, there have been several experimental studies which throw light (often of an indirect sort) upon psychoanalytical claims. A few pertinent results may be summarized as follows:

- (1) Experiments with animals and children lend some validity to the phenomenon of "regression," namely, that individuals when thwarted or blocked may revert to a simpler level of behavior.

- (2) Studies of the recall of early life experiences have, on the whole, failed to substantiate Freud's claim of infantile amnesia or the forgetting of the first few years of life.

- (3) There is experimental evidence that projection occurs when a person lacks insight and is unwilling, because of the pain or embarrassment

involved, to admit faults in himself. Projection describes the "mechanism" by which one "unconsciously" ascribes one's own motives to some other person or thing. "An unskilled workman blames his tools" is a commonplace illustration of projection.

(4) Dreams can be created experimentally and may reflect fears, anxieties, and various waking experiences. There is no evidence, however, that dreams are necessarily caused by "unconscious" conflicts.

(5) Careful experimental studies have shown that the frustration or blocking of a much desired goal often, but not necessarily, leads to aggressive behavior which is directed against the frustrating object or whatever represents it.

It will be clear from the above list that the experimental verification of psychoanalytical theories has not been particularly fruitful. Freud's ideas concerning motivation grew out of a study of the case histories of those of his patients who were suffering from various forms of neurosis. Some of Freud's "hunches" are difficult to test, while others are not susceptible to laboratory control. Whatever the future of psychoanalysis may be, it is fair to say that Freud's emphasis upon the rôle of the emotions in everyday life, his stress upon the importance of early experiences within the family, and his theories of "unconscious" motivation, have greatly stimulated thinking and controversy in scientific psychology. But it is as an innovator and originator rather than as an experimenter that scientific psychology is indebted to Freud.

3. FACTORS IN PERSONALITY

Many factors enter into and condition our personalities. Some of these are deep and permanent, others temporary and ephemeral. Consider, for example, the breadth and depth of a man's emotional responses; the quickness with which he loses his temper, is depressed or frightened. Insofar as these characteristics are basically physical and physiological, they are inborn and inherited though their manifestations are determined by training. On the other hand, good breeding and the social graces of tactfulness, courtesy and deference are habits which, to a large extent, the growing child acquires much as he learns arithmetic. Psychologists have enumerated various factors which they believe to be important in shaping personality. Among these we may select, as of special significance, appearance and physique, intelligence, temperament, and social adaptability.

(1) Appearance and physique are undoubtedly influential in determining characteristic behavior. The problem of social adjustment of the handsome, well-proportioned child, for example, is very different from that

of the ill-favored, weak and puny child. When a boy is teased or badgered at home or in school, such relatively minor handicaps as wearing glasses or a brace on the teeth or having large hands and feet assume exaggerated importance and lead to real unhappiness. Insecure children become super-sensitive, hypercritical and arrogant, or timid and apologetic, depending upon the strength of other compensating and balancing personality traits. Psychological adjustment often poses a serious problem for those who are deformed or crippled, or who because of a weak heart or other organic defect, cannot take part in normal activities. In such individuals dejection and feelings of inadequacy may be intensive and acute.

(2) A child's mental ability may greatly affect his social and emotional behavior. The slow and backward student is beaten before he starts, and his behavior often has the character of desperation. Intense feelings of insecurity—expressing themselves in overactivity or in withdrawal—may develop in the mentally retarded child. And the adult who is dull—and knows it—often shows the same typically resentful behavior. He may be morose and negative or rude and defiant. Education and training adapted to an individual's needs is good mental hygiene in such cases. The charge that the colleges are ruining a lot of potentially good janitors and truckdrivers is, perhaps, too strong, but it is certainly true that there are many youngsters in college who would be better off elsewhere.

(3) Temperament refers to an individual's characteristic emotional responses—the breadth and strength of his reactions and the degree of his self-control. The timid person who is easily cowed and browbeaten; the highly irritable man who flares up at the slightest provocation; the erratic person who is at one time dominating and superior and at another pleasant and agreeable—all of these have personality difficulties of a temperamental sort. People who lack emotional stability and self-control usually have trouble in making new friends and in keeping old ones.

(4) Social adaptability includes those responses, attitudes and opinions, intellectual, temperamental, and moral, which are exhibited in social participation and everyday relationships. Social adaptability depends to a large degree upon the factors listed above; and to a still greater degree than is true of most of these factors is dependent upon training and environment. This trait is given a separate heading in order to stress the fundamental importance for personality of behavior in a social setting.

4. IMPRESSIONISTIC METHODS OF STUDYING PERSONALITY

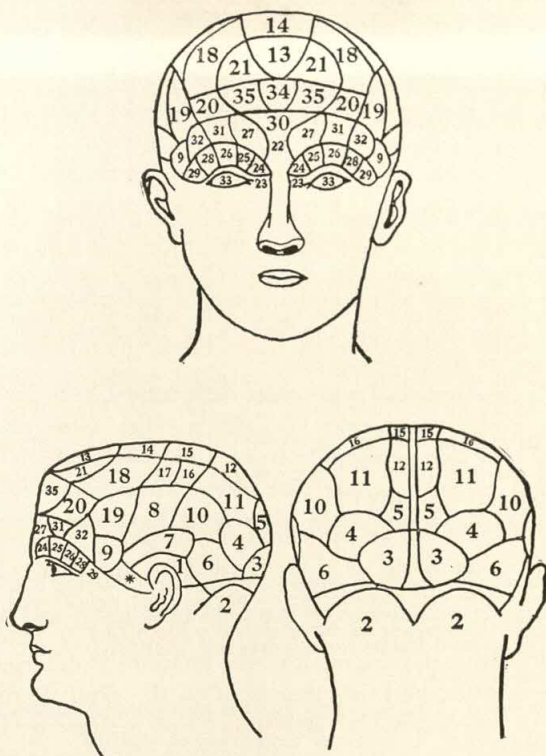
Attempts to discover and define personality traits may be classified in a general way as *impressionistic* or *experimental*. Impressionistic methods

are characterized by intuition rather than by logic. The claims made by their supporters are often extravagant and the results, compared with these claims, exceedingly meager. Illustrative of impressionistic methods of revealing character and personality traits are (1) systems of physiognomy; (2) various schemes dealing with body build or glandular makeup; (3) divination methods as exemplified in astrology, palmistry, and graphology.

(1) *Physiognomy*. This pseudo-science claims to reveal character and personality traits from an examination of bodily—and especially facial—characteristics. The conformation, size and shape of the nose, ears, chin, mouth and face; the texture of the skin; hair- and eye-color—all of these are taken to be indices of personality traits. Physiognomy reached its most ambitious modern development in vocational guidance. Blackford and Newcomb, for example, in their book "The Job, The Man, The Boss" published in 1924, give detailed directions for interpreting in personality terms the meaning of various facial indices, quality of the voice, manner of walking, style of clothing, and the like.

Phrenology, the best known branch of physiognomy, seeks to read personality traits from the contours of the skull. The "mind," according to the phrenologist, is composed of faculties—intellectual, emotional, and moral; and these hypothetical agencies have definite locations in "centers" within the brain. The relative development of the faculties is determined by observing the depressions and protuberances on the skull. Phrenology was introduced by two Germans, Gall and Spurzheim, in the early 1800's. Their system, which is essentially the one still in use today, marks off thirty-five areas on the skull. These anatomical divisions locate the underlying faculties to which are given such high sounding names as amativeness, benevolence, calculation, causality, and the like. (See Fig. 32.)

Physiognomists and phrenologists still appear at fairs and carnivals where they can practise their art upon the gullible. But none of these systems of personality analysis has any scientific standing whatever. Modern physiology has completely exploded the phrenological faculties. The supposed "organs" of reverence, conscientiousness, and hope, for instance, which are located on the roof of the skull, lie directly above the motor areas which control the movements of the hands and feet. Philoprogenitiveness and inhabitiveness lie just above the brain areas concerned with vision. Every controlled experimental test of the theories of physiognomy and phrenology has refuted their claims. There is no evidence of a direct relationship between either intellect or character and facial contours, skin texture, hair color, or other physical attributes. In the continued belief in these hypothetical relationships one often detects a high degree of wishful thinking. For instance, the strong chin, the beetling brow, the aquiline nose, the high forehead, are judged to be indicative of desirable personality traits insofar as they are possessed by members of our own racial or



AFFECTIVE FACULTIES

- PROPENSITIES
- ? Desire to live
 - * Alimentiveness
 - 1 Destructiveness
 - 2 Amativeness
 - 3 Philoprogenitiveness
 - 4 Adhesiveness
 - 5 Inhabitiveness
 - 6 Combativeness
 - 7 Secretiveness
 - 8 Acquisitiveness
 - 9 Constructiveness

- SENTIMENTS
- 10 Cautiousness
 - 11 Approbativeness
 - 12 Self-Esteem
 - 13 Benevolence
 - 14 Reverence
 - 15 Firmness
 - 16 Conscientiousness
 - 17 Hope
 - 18 Marvelousness
 - 19 Ideality
 - 20 Mirthfulness
 - 21 Imitation

INTELLECTUAL FACULTIES

- | PERCEPTIVE | REFLECTIVE |
|--------------------------|---------------|
| 22 Individuality | 34 Comparison |
| 23 Configuration | 35 Causality |
| 24 Size | |
| 25 Weight and Resistance | |
| 26 Coloring | |
| 27 Locality | |
| 28 Order | |
| 29 Calculation | |
| 30 Eventuality | |
| 31 Time | |
| 32 Tune | |
| 33 Language | |

FIG. 32
THE "POWERS AND ORGANS OF THE MIND"
(After Spurzheim, 1834.)

social group. For reviews of the experimental evidence on phrenology and physiognomy, see references (16) and (35).

{2) *Type theories.* From the time of Lombroso to the present, there

have been many attempts to relate body form and proportion to temperamental and personality traits. One of the most ambitious of these schemes was that of the German psychiatrist Kretschmer, who proposed a system of parallel physical and psychological types. On the basis of body build, Kretschmer classified mankind into three major groups, (1) the tall, slender and (usually) weak; (2) the strong, well proportioned and vigorous; (3) the short, round and fat. According to Kretschmer, the first group tends to be reclusive and self-contained and to exhibit shy, timid and unsociable behavior; the third type is inclined toward emotional instability and is subject to uncontrolled and boisterous outbursts which alternate with moods of melancholy and depression. In the main, the middle type exhibits balanced or normal behavior.

Kretschmer's constitutional types were determined from numerous body measurements, and his own work revealed a rather close association between physical type and personality traits. Later experimental work, however, has failed to substantiate the close parallelism which Kretschmer claimed, and in consequence his results have not been widely accepted.

The most thorough modern attack upon the relationship of physique and temperament is represented by the work of Sheldon and his collaborators. Sheldon proposed three basic physical components or types, which are described as follows.

(a) *Endomorphy*: A high rating in this component indicates a fat, soft person inclined toward "tubbiness" and short in stature.

(b) *Mesomorphy*: A man rated high for mesomorphy is a "bone and muscle man," well proportioned and strong.

(c) *Ectomorphy*: A high rating in this component indicates that the subject has long arms and legs, is thin or even skinny, with weak, flabby muscles.

A person is rated in each of these three basic types separately upon a seven-point scale. Characteristic body build (or somatotype) is defined by three numbers which together represent an individual's standing in endomorphy, mesomorphy, and ectomorphy. One represents the lowest, seven the highest standing in a given component. Thus, 7-1-1 denotes typically a soft, fat person; 1-1-7 a tall, skinny one. Two men marked 2-6-2 and 3-5-2 are both mesomorphic, the first being more nearly representative of this type than the second. Figure 33 illustrates the three typical somatotypes, and shows a fourth or balanced type. Sheldon's three physical categories are closely related to Kretschmer's constitutional types. But by describing an individual in terms of three components, each along a scale

of seven degrees, Sheldon avoided the all-or-none character of the earlier typology.

In a study of 200 college men, Sheldon has reported a strong association between ectomorphy (tall-thin) and such introverted personality characteristics as timidity, liking to be alone, emotional restraint, and nervous

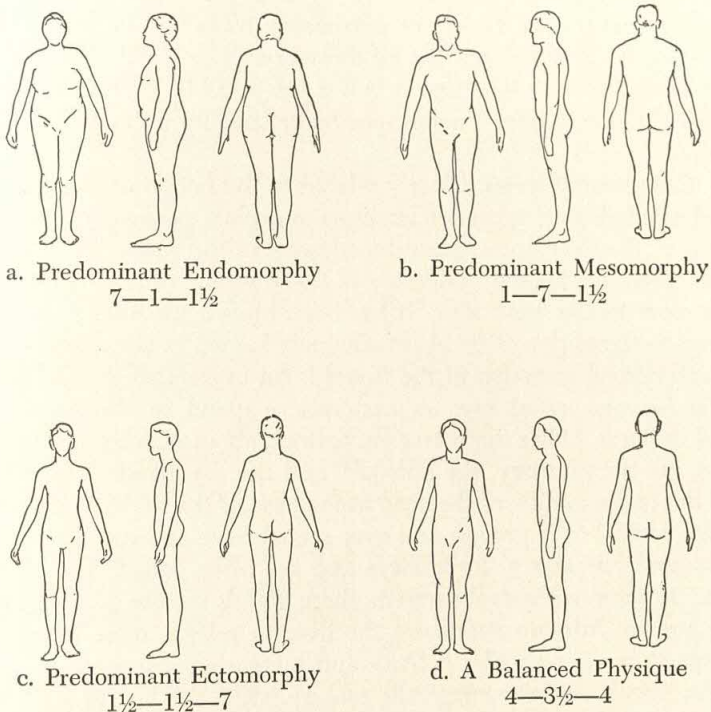


FIG. 33
HUMAN SOMATOTYPES

(From Sheldon and Stevens, *The Varieties of Temperament*, 1942. Harper & Brothers.)

instability; between endomorphy (short-fat) and such extroverted personality traits as love of comfort, conviviality, complacency, and free emotional expression. The intermediate type, or mesomorph, was inclined to be assertive, energetic, dominating, and often callous. Personality traits were determined from a large number of ratings based upon extensive interviews and observation.

Various investigators have checked Sheldon's results and have come out with conclusions which are not as striking as those reported by him. Cor-

relations of somatotype with objective tests and personality inventories, for example, have revealed only a moderate association between physique and temperament. Various criticisms of Sheldon's procedures have also been offered. Perhaps a certain correspondence between physique and personality is to be expected—strong men, for example, can afford to be more complacent, say, or more assertive, than weak men. But it is a safe conclusion that a wide variety of personality traits will be found in men of the same body type; and that all three body types will often show the same personality trait or traits. It is not yet possible, therefore, from his body build alone to infer a man's specific personality traits with any great precision.

(3) *Glandular theories.* Closely related to the notion of parallel physical and psychological types are attempts to explain personality differences in terms of the development and functioning of the endocrine or ductless glands. The best known endocrine is the thyroid, situated in the neck region close to the wind pipe. It has been known for some time that in children the condition of feeble-mindedness known as cretinism is caused by a sub-normal secretion of the thyroid. An overactive thyroid leads to general nervous irritability; an underactive gland to sluggishness and mental dullness. Other important endocrines, about which considerable is known, are the pituitary, the adrenals, and the sex glands. The pituitary gland lies in the middle of the head at the base of the brain. It is known to exercise control over growth and over sexual development. The adrenals are situated just above the kidneys and are often called "the glands of action." In intense physical exertion, these glands secrete adrenin into the blood stream. Adrenin stimulates the liver to release more blood sugar, thus speeding up muscular activity and increasing resistance to fatigue. The sex glands are physiologically mature when the boy or girl reaches puberty. Secretions from these glands exercise control over primary sexual development, and over secondary sex characteristics as well. These latter include voice changes, body form, muscular growth, hair distribution, and other attributes which are distinctively masculine or feminine.

Enthusiastic endocrinologists have attempted to link up definite personality traits to under- or overactivity of one or more of the ductless glands. Shyness, timidity, and feelings of inferiority are said to result from deficient thyroid secretion; aggressiveness, expansiveness, feelings of superiority, from an overactive gland. "Moral and intellectual inferiority," compulsions, and pathological lying have been attributed to lowered activity of the pituitary; pugnacious and aggressive behavior to overactivity of the adrenals and timorous and vacillating conduct to underactivity of these same glands. These speculations, while interesting, are not well established experimentally, and except in abnormal or extreme cases are probably

much exaggerated. At the same time, much of scientific value bearing upon the physiological basis of personality has resulted from investigations in this field.

(4) *Divination methods of judging personality traits.* Pseudo-scientific systems like palmistry and astrology rank with physiognomy and phrenology, and are on a decidedly lower level than are the type and glandular theories just described. Palmistry and astrology may be dismissed at once from serious consideration. Both of these represent exploded and discredited efforts to find a royal road to the understanding of human nature, and both now belong in the class of parlor amusements. If palmists or astrologers could produce evidence of high correlation between the lines in the hand, or being born under a certain "sign," and the possession of definite personality traits, some credence might be given to their claims. But quantitative studies will never be made by practitioners of these "sciences" as long as people are willing to spend money without asking for evidence.

Graphology has far more to offer as a means of inferring personality traits than the schemes just discussed. Handwriting is overt expressive behavior, not a physical "sign" or static index. From a person's handwriting we can usually tell whether the writer is a man or a woman, is old or young. We can infer, too, from the handwriting something about a person's education and state of health. Handwriting is so susceptible to practice and to training, however, that it is doubtful whether any specific information regarding personality traits can be obtained from it alone. Hull and Montgomery studied the relation between upward and sloping lines, heavy and fine lines, crossing of *t*'s, method of forming *m*'s, *n*'s, *o*'s, and the like, and such traits as ambition, timidity, and forcefulness, which, according to graphologists, are expressed by these characters. Results were decidedly negative. Allport's and Vernon's studies, as well as those of Wolff, have been somewhat more favorable to graphology. Handwriting specimens can be successfully matched against photographs, voice records, and character sketches. But the agreement is not close enough to permit prediction except in very general terms.

5. WHY IMPRESSIONISTIC METHODS HAVE VOGUE

Why do systems of personality and character analysis flourish when there is no substance to their claims? Several answers may be offered to this question. In the *first* place, many people possess little or no scientific training, and hence have no conception of what is meant by a rational world, in which events have causes and things do not happen by magic. Such people make up the naïve, gullible and mentally immature folk who

search for Fountains of Youth, universal panaceas, talismans, and cure-alls; who join esoteric cults and indulge in all sorts of emotional vagaries. *Secondly*, there is just enough truth in some of these schemes—physiognomy, for instance—to make them plausible. Much can obviously be inferred or shrewdly guessed about a person's early training, cultural background, habits and education, from his appearance, speech, mannerisms, occupation and interests. The effects of dissipation and disease, of a friendly or a crabbed disposition appear in the face, the conversation, the body carriage, and the gait. One can usually tell from a man's hands whether he is a bookkeeper or a mechanic; from his skin whether he is a rancher or a clerk; from his physique whether he is an athlete or a weakling. Man is a civilized creature living in a social environment, and as such reflects in his behavior and bearing the experiences through which he has passed. Some of the numerous physiognomic signs will occasionally, by chance alone, be associated with almost every personality trait. But this is quite different from saying that certain lines in the hand indicate longevity; that a high forehead means intellect; that a bump on the skull indicates "love of approbation"; that being born in March rather than in September makes the difference between being a successful or an unsuccessful lawyer. None of these "signs" is related to or expressive of personality in the same way that a student's performance on an aptitude test is indicative of his ability to succeed in engineering courses.

Still a *third* reason for the continued popularity of impressionistic methods grows out of the inveterate human tendency to classify. People are judged wholesale as good or bad, moral or immoral, sneaky or forthright, though rarely if ever does one deserve such an unqualified designation. Classification is a lazy way of avoiding analysis and of persuading ourselves that we have made an intelligent decision. Hasty and unwarranted judgments grow out of the tendency (a) to generalize broadly and with disregard of negative cases; (b) to reason by analogy; and (c) to mistake sequence for causal connections. Thus it is "common knowledge" that red-headed people are quick-tempered; that blondes are fickle; that long-fingered people are artistic; that pretty girls are empty headed. A cannon makes more noise than a high-powered rifle, and by analogy the man with a loud voice and energetic manner is more "forceful" than is a quiet, unassuming one. Because a friend was told by a palmist that she would marry a tall, blond man—and did—fortune telling is elevated to the dignity of an exact science.

The best antidote for impressionistic methods is the rigorous experimental approach. When a well-controlled experiment yields no evidence in favor of a given proposal, the hypothesis must be dropped outright—no matter how plausible or desirable—or altered and again subjected to test.

6. EXPERIMENTAL METHODS IN THE STUDY OF PERSONALITY

Experimental techniques used in the study of personality include the rating scale, the questionnaire, the objective tests, and various projective techniques. In many instances, perhaps in most, all of these methods will find some application in a given investigation. They are described in this section.

(1) *The rating scale.* It is customary in moral teaching to make a distinction between character and reputation. Character, we are told, is what we really are, and reputation is what people think we are. This pious contrast is perhaps more wishful than accurate. Character is reflected in reputation; and often for better or worse we are essentially what people say we are. Rating scales are devices for ranking or rating people with respect to their possession of personality and intellectual traits, skills, occupational efficiency, likes and dislikes, attitudes and interests. Supervisors rate clerks for appearance, intelligence, selling ability, cheerfulness, patience, and tact. Teachers rate pupils for ability to learn, study habits, and various behavior traits. An illustration of a graphic rating scale planned for use with students is given in Figure 34. The rater places a check somewhere along the line to indicate the degree to which the student—in his judgment—possesses the given trait. Descriptive phrases entered below the lines are for the guidance of the rater.

The average individual should be rated somewhere in the middle of a scale line. Raters often err on the side of leniency, however, and so tend to place more people above than below the middle point. Training and careful instruction are sometimes necessary to correct this tendency. Another difficulty with rating scales has been called "halo effect." Halo shows itself in the tendency of a judge to place a person of whom he has a good opinion high in every trait. To avoid halo, every subject should be rated on each trait in order rather than each subject separately rated on all traits. Makers of rating scales are careful to select behavior traits which are objective and readily defined; which are not too broad or too specific; and which can be observed in action rather than inferred from conversation or hearsay. It is easier to rate a filing clerk for accuracy of work, or a teacher for cheerfulness in the classroom, than to rate either for loyalty, sincerity, or depth of religious feeling. Psychologists do not put much confidence in a single rating, but the consensus of a group of competent judges can be an exceedingly useful index. In many schools, pupils are rated on a variety of personality traits year after year. A comparison of these records shows the direction in which the child is developing as well as his level of maturity at a given time.

Name of student

A. How are you and others affected by his appearance and manner?

Avoided by others	Tolerated by others	Liked by others	Well liked by others	Sought by others	No opportunity to observe

Please record here instances that support your judgment.

B. Does he need constant producing or does he go ahead with his work without being told?

Needs much producing in doing ordinary assignments	Needs occasional producing	Does ordinary assignments on his own accord	Completes suggested supplementary work	Seeks and acts for himself in additional tasks

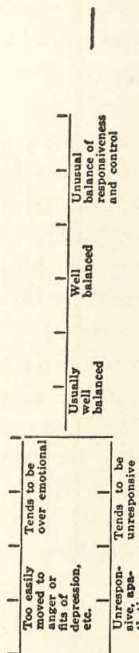
Please record here instances that support your judgment.

C. Does he get others to do what he wishes?

Probably unable to lead his fellows	Lets others take lead	Sometimes leads in minor affairs	Sometimes leads in important affairs	Displays leadership ability to lead his fellows; makes things go

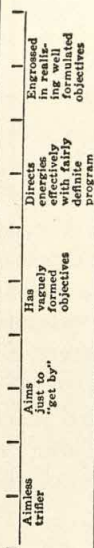
Please record here instances that support your judgment.

D. How does he control his emotions?



Please record here instances that support your judgment.

E. Has he a program with definite purposes in terms of which he distributes his time and energy?



Please record here instances that support your judgment.

FIG. 34
GRAPHIC RATING SCALE
(From Bradshaw, Arch. Psych., No. 119, 1930, pp. 44-45.)

(2) *The questionnaire.* A questionnaire presents a set of questions or statements to be answered or marked by the subject himself or by someone who knows him well. A questionnaire, unlike a test, is not scored in terms of time or errors or amount done. Instead, it is a systematic report which serves to identify the subject as like or unlike other individuals of the same sex, age level, social or occupational background. The questionnaire has been used by psychologists to obtain data upon (1) personal difficulties or troubles—fears, worries, feelings of inadequacy, and the like; (2) attitudes and beliefs; (3) interests in intellectual, mechanical, or social activities, as well as in sports, vocations, cultural and practical pursuits. Answers to a questionnaire may be categorical—given simply as yes or no—or varying degrees of assent or dissent may be permitted by providing several alternate replies. It will be profitable to consider the various usages of the questionnaire in some greater detail.

(a) The personal data sheets or neurotic inventories are attempts to get information concerning a person's fears, worries, social or personal difficulties. Items on these questionnaires can be classified roughly as (1) primarily physical symptoms, such as pain, fatigue, incoördination; (2) primarily emotional symptoms, anxieties, fantasies, depressions, sleep disturbance; (3) personal reactions, to liquor, drugs, sex; (4) mental symptoms, such as vacillation, compulsion, queer feelings about oneself or other people. Sample questions are:

	Yes	No
Do you usually feel strong and well?	—	—
Does your heart ever thump in your ears so you cannot sleep?	—	—
Do you have a great many bad headaches?	—	—
Does liquor make you quarrelsome?	—	—
Are you troubled with fears of being crushed in a crowd?	—	—
Does your mind wander badly so that you lose track of what you are doing?	—	—

Motivation is especially important in questionnaire data, as there is usually no sure guarantee that the person is telling the truth. Several ingenious techniques have been developed for detecting dishonest answers. Presumably answers will be trustworthy when the respondent feels that an honest reply will be of advantage to himself. When individual reports are not important and only group tendencies are desired, anonymity may be resorted to as a means of insuring more nearly objective replies.

Sometimes a questionnaire is so constructed that the main purpose of the inquiry is concealed from the subject, who then gives desired information unwittingly. One way of doing this is to give the subject a choice as between two types of reading matter (popular science—romantic stories), two sports (fishing—football), several occupations (actor—engineer—

social worker). The respondent is instructed to select what are, for him, the most desirable—or least undesirable—of the choices presented. This has been called the method of “forced choices.” Dominant personality trends and attitudes are often revealed by examining and relating the individual’s choices. It is hard for most people to outguess the test or divine the significance of their choices.

(b) Attitudes often reflect basic personality traits, so that a knowledge of the consistency and generalized character of a person’s feelings toward controversial issues is often valuable. One comprehensive set of attitude scales undertakes to measure characteristic dispositions toward various social, economic, and ethical problems—toward the church, the movies, the Negro, war, communism, and the like. Items from a questionnaire designed to measure attitude toward the movies will provide samples:

Movies are the most important cause of crime.

Movies are good clean entertainment.

Going to a movie is a foolish way to spend your money.

I like the movies as they are because I go to be entertained, not educated.

A person’s score depends upon the character of the statements which he accepts or endorses. Scale values are determined for each item from an analysis of the judgment expressed by an experimental group. The “weight” or numerical value of a statement is a measure of the extent to which the statement favors, or does not favor, the institution or issue under consideration. Low scale values denote favorable, high scale values unfavorable attitudes. In the examples listed above, the second statement has a lower (more favorable) weight than the third, and a much lower scale weight than the first.

Attitude scales have been used to compare groups of different background and training; to measure the effects of certain types of instruction, and of various kinds of propaganda; to study the relations of attitudes to such factors as intelligence, special aptitude, and socio-economic status.

(c) A third approach to the questionnaire study of personality differences is through investigation of the interest reported by groups vocationally or culturally dissimilar. The vocational interest blank designed by E. K. Strong is a comprehensive survey of occupational interest, of attitudes, and of everyday likes and dislikes. The man who marks this vocational interest blank expresses his preferences for 100 occupations listed in order, his likes and dislikes for various school subjects, for amusements and recreations, and for many commonplace activities such as writing a letter or decorating a room. The subject must also tell how various peculiarities of people affect him, his attitude toward different kinds of work (indoor *versus* outdoor), as well as his judgment concerning his own abilities

and personal traits. Different scoring keys permit the blank to be tallied separately for interests of a lawyer, architect, teacher, printer, and many other pursuits. An individual's score on the scale of "lawyer interests," say, will identify him as having the interests of a lawyer, as probably having such interests, or as definitely not having such interests. Information of this sort is especially valuable in vocational guidance.

In another useful interest questionnaire (the Kuder Preference Record) the individual indicates his preferences and interests for a wide range of activities. Choices are classified as falling within nine interest areas: mechanical, computational, scientific, persuasive, artistic, literary, musical, social service, clerical. Scores within each area are represented on an "interest profile" in order to show relative strengths and weaknesses. From his profile, a man can usually tell within what "family" of occupations his dominant interests place him. One man, for example, may show strong interest in mechanical and computational areas, weak interest in persuasive, literary, and musical areas; another man strong interest in social service, persuasive, and artistic areas, and weak interest in clerical and computational areas. As might be expected, the relationships among interest scores in related fields (artistic and literary, for example) are quite high.

(3) *Personality tests.* Many personality tests are really questionnaires which resemble objective tests in construction and administration. An interesting example is the George Washington Social Intelligence Test. In taking this "test," the respondent is called upon to give a judgment involving a social situation by marking one of several alternatives; to infer the "mental state" of a speaker from a list of exclamations and expressions; to decide whether certain descriptions of human behavior depicted in various situations are true or false; to tell which of several possible replies to a given question is the most humorous. Only one section of the test, that dealing with memory for names and faces, presents a task which is scorable in terms of performance. In the other four parts of the test, answers are scored in accordance with standards set up by a consensus of experts.

Performance tests are often used as measures of persistence and flexibility, of muscular effort, of estimation of distance and direction. Even tests of arithmetic and school information have served as indicators of changes in attitude or motivation (see p. 189). Various physiological indices have sometimes been employed as measures of emotional state and of muscle tension. Typical examples are pulse and breathing rate, blood pressure, and the psychogalvanic reflex.

The hypothesis that people reveal their thoughts, feelings and attitudes in their spontaneous creations or unguarded responses is the basis

of the so-called "projective techniques" often used in systematic studies of personality. In projective tests there are no right and wrong answers, an almost unlimited range of responses being permitted. The hope is that the respondent will "project" into his performances persistent personal problems, worries, fears, chronic emotional attitudes, which he is unwilling or unable to report. Several sorts of stimuli having more or less "structure" have been used to elicit projective effects. Among such materials are ink blots, cloud pictures, free drawings, the arrangement of toys, dramatic play, photographs and pictures, and free completion of sentences. The more nebulous and unstructured the stimulus, the greater the probability of free imaginative flights. Several projective techniques are illustrated in the studies reported on pages 195-196.

7. SYSTEMATIC EXPERIMENTAL STUDIES OF PERSONALITY TRAITS

Studies of personality often suffer from a lack of theoretical orientation, a failure properly to define variables, inadequate sampling, and crude experimental techniques. The studies summarized in this section are representative of the best experimental work in this difficult field.

(1) *The Hartshorne-May Studies in Deceit*. These investigations are careful and painstaking studies of the character trait of honesty. Hartshorne and May set out to discover what variables accompany and perhaps condition honest behavior; and whether honest behavior is a general characteristic or is, instead, a group of loosely related habits to which we apply a common descriptive term. Put differently, is an "honest man" *always* honest, or is he honest most of the time, or now and then.

Several thousand children mostly in grades 5 to 8 constituted the experimental sample. Honesty, as shown in resistance to cheating, was measured in a variety of ways. Test situations were set up in the classroom, on the athletic field, and in social situations out of school (parlor games), in all of which cheating or deceptive behavior was at least feasible. Scores achieved in classroom tests of arithmetic, word knowledge, reading and spelling administered under close supervision, were compared with scores earned under "dishonest" conditions, that is, under conditions in which cheating could readily take place. Sometimes children were allowed to correct their own test papers, the scores under honest conditions having first been recorded without the child's knowledge; or, the "honest maximum" for a test having been determined, scores above this level were taken as *prima facie* evidence of deception.

In out-of-school situations, "improbable achievement" was often taken as indicative of deceit. To illustrate, an athletic contest was arranged

in which chinning the bar, broad jumping, breathing into a spirometer, and strength of grip tests were featured. It has long been known that fatigue sets in rapidly after the first few trials with the hand dynamometer—an instrument which measures hand strength in terms of maximum squeeze. Accordingly, a secret record was made of the first three trials, and the child then encouraged to practise on his own, and later to report his achievement. Any large increase beyond the initial scores is most unlikely, and hence is evidence of cheating. The same is true of breathing capacity and of the other motor tests: marked improvement beyond a certain point rarely if ever occurs. The “improbable achievement” technique was introduced into several of the parlor games. It is unlikely that a blindfolded child will pin the tail unerringly on a donkey when told to approach the figure from some distance. If the child quickly succeeds, he may be suspected of having peeked under the eye bandage, of having followed the lines on the floor, or of having made unfair use of other visual cues. In another game, objects of various sorts including a number of dimes, were passed out to children who were told to deposit their “gifts” in a box in an adjoining room. As no *apparent* check was made on whether the child did as he was told, retention of the dime was an attractive possibility.

Hartshorne's and May's study yielded many suggestive results. Older children cheated more often than younger; the emotionally unstable more often than the normal; retarded children more often than unretarded. There was a substantial correlation between honest behavior and intelligence, and between honesty and social and economic status. Attendance at Sunday School and religious instruction apparently had little influence upon honesty when socio-economic level was ruled out. The morale of the classroom proved to be quite important. Under some teachers, for instance, practically none of the pupils cheated; under other teachers practically all of them did.

From the very low relations among their tests, Hartshorne and May concluded that there is no *general trait* of honesty. This is perhaps their most important single conclusion, and it had far reaching implications. If honesty is reducible to a congeries of fairly specific habits, there is little meaning to the designation “an honest man.” The best training in honest behavior which the teacher can provide under these circumstances would consist in the cultivation of specific habits of honesty: teaching honesty as an ideal would apparently be of little value. This conclusion of Hartshorne and May has been criticized on several counts. Teachers and child psychologists recognize that children do not have the same standards of honesty as adults, and are not motivated by the same considerations. School, for example, is often a game in which the pupil is

pitted against the teacher; and cheating under such conditions is regarded as a clever and praiseworthy accomplishment, in no sense a violation of the rules of good conduct. Rivalry, competition, desire to please, as well as varying degrees of interest and effort, undoubtedly entered into Hartshorne's and May's tests, and such factors may well have outweighed the "honesty component" which the tests were designed to assay. Low relationships among the tests might reflect, then, a multiplicity of motives—some working at cross purposes—rather than the absence of a single dominant motive. Common observation and the consistency found in studies of attitudes make it appear likely that honesty is more generalized than this study would indicate. Certainly everyday transactions are predicated on the assumption that honest behavior is generally predictable: that in "honest people" one sort of behavior is far more likely than the opposite.

(2) *The Allport-Vernon Studies of Expressive Movements.* These authors were interested in the extent to which expressive movements are consistent or harmonious within the same individual, and in whether such movements can be related to measured personality variables. Expressive movements were defined as characteristic motor acts such as handwriting, reading, drawing and counting, tempo of walking and strolling, accuracy in the estimation of sizes, distances and angles, strength of grip, speed of tapping. In addition to objective tests, ratings were obtained as to fluency of speech, intensity of voice, rapidity, smoothness and rhythm in gesticulation and movement.

Repetition of the test situations showed considerable consistency in the characteristic ways in which an individual expressed himself. Moreover, the relationships among the different tests were positive and rather marked, while the tests were closely enough related to the ratings to indicate surprising agreement between these judgments and actual performance. Knowing the speed, tempo and rhythm of some of an individual's activities, one can predict quite well what characteristics other expressive acts will have. Even when consistency is not so evident, Allport and Vernon believe there may still be an underlying agreement or "congruence." That is, the same fundamental motive often runs through activities which on the surface appear to be contradictory.

In their studies of handwriting, Allport and Vernon found that skilled graphologists could match handwriting specimens and thumbnail sketches of a person considerably better than chance would allow. These experiments indicated that handwriting is more expressive and more revealing of personality than many of us have believed, though nothing like so accurate an index as professional graphologists have claimed. Handwriting may often be indicative of aggressiveness, timidity, impulsiveness or

sensitiveness, but if so it must be judged by experts. One's writing is not an index which "he who runs may read."

(3) *The Rorschach Test*. The Rorschach test was devised by the Swiss psychiatrist, Hermann Rorschach. Although usually classified as a projective test, in reality the Rorschach is a controlled test-interview. Test stimuli consist of ten irregular ink blots which stand out against a white background. Five of the blots are in black and grey and five contain color—red, yellow, blue, and green. In the administration of the test, a subject is given one card at a time; he is allowed to look at the blot from any direction and for as long as he likes; and is then told to give all the associations and meanings which are suggested to him by the ink blot. To illustrate, a patient suffering from schizophrenia (shut-in, extremely introverted condition) gave the following responses to Card #1: "It resembles a butterfly with some sort of feelers on it and protruding eyes; a lizard's tail; a double body . . . made of two materials . . . the outside is transparent so you can see the inner; the open parts don't fit in . . . yes, they do; and those are wings, they are attached in three places." Using a detailed scoring procedure, the examiner first classifies the subject's responses into such categories as form, quality, detail, wholeness, originality, color and movement. Then from an analysis of his replies the dominant personality trends of the subject are inferred. Common, unoriginal responses, and those involving animals are interpreted to mean low intelligence; whole responses, good form and movement, high intelligence; detailed responses are taken to be indicative of a reclusive personality, while color responses mean emotional instability.

Several criticisms have been leveled against the Rorschach method. It has been argued that the test lacks objectivity in administration and scoring; that norms are scanty or not existent; that interpretation is intuitive and without experimental verification. In spite of these valid objections, the diagnosis of mental disease by the Rorschach is often strikingly confirmed by clinical findings. Insight, too, has been afforded into the personality difficulties of delinquents and feeble-minded. While too much has been claimed for the Rorschach by its enthusiasts, in the hands of a trained examiner it is undoubtedly a valuable instrument.

(4) *The Terman-Miles Studies of Masculinity and Femininity*. This study is a thorough investigation of sex differences in personality trends, attitudes, and interests. By means of a test of over 400 items, these psychologists obtained a masculinity-femininity index which locates an individual along a hypothetical continuum ranging from intensely masculine to intensely feminine. The test has seven parts dealing respectively with (1) word association, (2) ink blot association, (3) information, (4) emotional and ethical attitudes, (5) interests, likes and dislikes, (6) opinion,

mostly about people, and (7) introversion-extroversion. A response is scored + if it is typically masculine, — if it is typically feminine. Thus, to associate *meat* with *tender* is masculine, to associate *kind* or *loving* with *tender* is feminine; reading detective stories, liking chemistry, wanting to be a forest ranger are masculine interests, while reading poetry, liking dramatics, and wanting to be a social worker are feminine interests.

The "diagnostic value" of the items—whether indicative of masculine or of feminine attitudes—was determined from a compilation and comparison of responses obtained from many male and female groups. Hundreds of subjects, children and adults, in school and out, athletes, various occupational groups, delinquents, and homosexuals were tested in the process of standardizing the questionnaire.

According to the results of these authors, males are in general more extroverted than females, less sentimental, more interested in tools, machines, and business affairs. Females are more introverted than males, more sympathetic and more emotional; more interested in occupations requiring personal service. This contrast in masculine and feminine interests follows the traditional pattern and is certainly not surprising. More striking, perhaps, is the finding that scores made by men and women of culture and education differ very little. Also interesting is the fact that so-called typical male-female differences become more and more marked as we go down the social and economic scale. Validation of the scale is found in the strongly feminine scores made by male homosexuals.

(5) *Factorial Studies of Personality Organization*. A basic personality trait can be regarded as a relatively permanent disposition consistently revealed in a variety of characteristic behaviors. Since people are more or less cheerful, aggressive or austere, individuals can be arranged along a behavior continuum from cheerful to gloomy, aggressive to submissive, austere to amiable. Personality traits, therefore, are conceived to be bipolar and to vary along a dimension between two extremes. In the location and measurement of such personality dimensions, a correlational method called "factor analysis" is coming to be more and more widely used. If we know the intercorrelations among a set of personality tests, we can through factor analysis extract the common bond or bonds which run through the related variables. These statistically defined "factors" are then given meaning and body by reference to those tests in which they have the heaviest weights.

Several factor studies of personality have been carried out by the Guilfords. Those test items which appear most frequently in questionnaires designed to measure introversion-extroversion were correlated, analyzed and regrouped, so that a high degree of homogeneity was obtained in the personality variables finally chosen. Five bipolar factors

covering roughly the general dimension of introversion-extroversion were identified as follows:

Social introversion: A high score indicates shyness, tendency to keep in the background at social affairs; a low score means social-mindedness, enjoyment of company.

Thinking introversion: A high score indicates reflection, meditative thinking; a low score impulsiveness, outward looking.

Depression: A high score means frequency of blues, worry over events and possible misfortunes; a low score cheerful, optimistic disposition.

Cycloid tendency: A high score implies ups and downs of moods; a low score stable, emotional reactions.

Rhathymia: A high score means happy-go-lucky, carefree attitude; a low score caution, inhibition of impulses.

Several of these personality factors have been identified by other investigators. They are not necessarily independent variables, and some of them probably describe much the same behavior. It is difficult for an experimenter to decide just how homogeneous a personality dimension should be, and how independent of other dimensions, in order to be given status as a common trait.

Factorial methods have been applied to the analysis of behavior ratings of personality as well as to questionnaire items. By grouping together those behavior descriptions which are highly related and hence presumably indicative of the same personality dimension, Cattell has defined twelve factors which he calls "primary source traits of personality." These dimensions are:

Shut-in vs. outgoing responses

High vs. low mental ability

Emotional maturity vs. emotional immaturity

Ascendance vs. submission

Placidity vs. agitated behavior

Sensitiveness vs. toughmindedness

Cultured vs. boorish manner

Highly integrated vs. poorly integrated behavior

Coöperative, charitable vs. obstructive, withdrawn behavior

Vigorous vs. lethargic action

Hypersensitive vs. phlegmatic responses

Cheerful, realistic vs. obsessed, paranoid attitudes

Personality dimensions defined by factor analysis are of primary interest to psychologists because they represent generalizations which cut across

a wide sample of behavior. The existence of such statistically defined common traits is not opposed to the view that the configuration of traits *within* a given person is unique. (p. 171).

(6) *Murray's Explorations in Personality*. These comprehensive studies represent an attempt to obtain coherent pictures of individual personalities. Personality is analyzed in terms of *needs* which may be *viscerogenic* (hunger or thirst) or *psychogenic* (prestige seeking or insecurity). The psychogenic needs, though closely related to the viscerogenic, are conceived to stem principally from the social pressures in the environment. Behavior is analyzed in terms of these needs into various categories which are studied as separate dimensions and in terms of their interrelations.

On the whole, the approach in these studies is more intuitive and psychoanalytical than experimental, though many objective measurements were recorded. As the first step in the analysis, each subject was interviewed by the five members of a Diagnostic Council, each of whom rated him on forty traits. The subject wrote an autobiography, and answered three questionnaires covering personal data ("troubles"), his emotional and social attitudes and interests. His sex history was then recorded by a psychiatrist and he was put through some twenty tests and experiments. There were 51 subjects—all college students—about one fourth of whom gave thirty-six hours to the investigation.

The dominant needs of an individual—both the verbally expressed and the unexpressed (unconscious or latent)—were diagnosed from fantasies and free associations stimulated by various word lists, stories, poems, pictures. Some of these test situations may, for clarity, be described in more detail. In one test—called the esthetic appreciation test—the respondent was asked to pass upon the artistic merits of ten poems and ten pictures representing a variety of subject-matter. Fundamental attitudes as well as taste and insight were presumably revealed by these judgments. Another test measured the "level of aspiration" defined as the level of performance in a familiar task (for example, printing words with a handprint set) which the subject set for himself and attempted to reach. Tendency to repress unpleasant experiences was ascertained by seeing whether the subject forgot more readily those jigsaw puzzles at which he had failed than those at which he had succeeded.

Two test situations were especially novel and revealing of personality trends. In the first, the *thematic apperception test*, the subject was asked to interpret or tell a story about each of a set of pictures; in the second, a dramatic productions test, the subject was asked to arrange a group of small toy figures so as to depict a dramatic scene. Both tests proved to be useful methods of obtaining evidence regarding the subject's inner needs and in throwing light upon the environmental and social difficulties

facing him. One subject, for example, who had lived through persecution and other cruelties as a child, obviously "projected" this background into his replies. In one of the pictures presented to each subject the figure of a boy is seen huddled on the floor against a couch. His head is bowed on his right arm, and beside him on the floor is an object which resembles a revolver. Of this picture, our subject said:

Some trouble has occurred. Some one he loved has shot herself. Probably it is his mother. She may have done it out of poverty. He, being fairly grown up, sees the misery of it all and would like to shoot himself. But he is young and braces up after a while. For some time he lives in misery—the first few months thinking of death.

Because of its scope and detail, it is difficult to evaluate the work of Murray and his colleagues in brief space. It may be said in favor of the intuitive approach to personality analysis that the case study material offered is decidedly informative. Moreover, it seems likely that almost everyone would be benefited by such a searching personality analysis as Murray's techniques make possible. On the other hand, the method is time consuming, requires a highly trained personnel, is relatively unstandardized, and is difficult to repeat. Some validation has recently been provided by the application of these methods during World War II. Many of the projective and other techniques used in this study were adapted, altered, and expanded in order to provide assessments of the personalities of men for various types of military service. Systematic studies like these show the possibilities inherent in the psychoanalytical approach, and what may be expected if and when intuitive insights are brought under laboratory control.

(7) *Topological studies of personality.* Lewin and his colleagues have studied personality through an analysis of situations which provide conflict or dilemma behavior. Conflicts are described in terms of what is called "field theory" or topology. The field is the environment or life space in which the individual finds himself, and is represented geometrically as a blueprint or diagram. Forces (motives) acting upon the person from within or without, goal sought, routes to these goals, barriers and other obstructions, are represented graphically by lines and by various symbols in the behavior field.

Many social psychologists have found Lewin's diagrams useful in representing the interplay of motives among persons in a social situation. Lewin's use of elementary mathematical concepts to represent motives and their frustrations is perhaps more picturesque than useful, but many ingenious experiments have been carried out within the framework of field theory. The following are examples of some of the problems attacked and results obtained: (a) Under free play, children were evaluated by

observers for degree of "constructiveness." When interfered with or frustrated, these children resorted to a lower level of constructive play. (b) Success or failure in a given task caused a shift in the aspiration level (goal set) in other tasks. (c) Goal objectives or aspiration level was discovered to be a fairly general personality trait. (d) Preadolescent boys showed less aggressiveness and worked harder in a so-called "democratic atmosphere." In an autocratic situation, hostility was likely to appear, especially toward the leader.

BIBLIOGRAPHY

1. Allport, G. W. A test for ascendance-submission. *J. abnorm. (soc.) Psychol.*, 1928, 23, 118-136.
2. Allport, G. W., & Vernon, P. E. *Studies in expressive movement*. New York: The Macmillan Company, 1933, chapters 9, 10, 11.
3. Anastasi, A. Review of Sheldon, W. H. and Stevens, S. S. *The varieties of temperament*. *Psychol. Bull.*, 1943, 40, 146-149.
4. Beck, S. J. *Introduction to the Rorschach method*. New York: American Orthopsychiatric Association, 1937.
5. Blackford, K. M., & Newcomb, A. *The job, the man, the boss*. Garden City, New York: Doubleday, Page & Co., 1924.
6. Bradshaw, F. F. The American council on education rating scale. *Arch. Psychol.*, 1930, 18, No. 119.
7. Brill, A. A. (Ed.). *The basic writings of Sigmund Freud*. New York: Modern Library, 1938.
8. Cattell, R. B. *Description and measurement of personality*. Yonkers, New York: World Book Company, 1946.
9. Child, I. L. and Sheldon, W. H. The correlation between components of physique and scores on certain psychological tests. *Character and Personality*, 1941, 10, 23-24.
10. Dollard, J., Doob, L., et al. *Frustration and aggression*. New Haven: Yale University Press, 1939.
11. Downey, J. E. *The will-temperament and its testing*. Yonkers, New York: World Book Company, 1923.
12. Freud, S. *The problem of lay analysis*. New York: Brentano, 1927, p. 68.
13. Fryer, D. *The measurement of interests in relation to human adjustment*. New York: Henry Holt and Company, 1931, chapters 3, 4.
14. Garvey, C. R. Comparative body build of manic-depressive and schizophrenic patients. *Psychol. Bull.*, 1933, 30, 567-568.
15. Gilliland, A. R. A revision and some results with the Moore-Gilliland aggressiveness test. *J. appl. Psychol.*, 1926, 10, 143-150.
16. Griffiths, C. H. *Fundamentals of vocational psychology*. New York: The Macmillan Company, 1924, chapters 3-4.
17. Guilford, J. P. and Guilford, R. B. An analysis of the factors in a typical test of introversion-extroversion. *J. abn. and soc. Psychol.*, 1934, 28, 377-399.
18. Hartshorne, H., & May, M. A. *Studies in deceit*. New York: The Macmillan Company, 1928.
19. Heidbreder, E. Measuring introversion and extroversion. *J. abnorm. (soc.) Psychol.*, 1926, 21, 120-134.

20. Hoskins, R. G. *The tides of life*. New York: W. W. Norton and Company, 1933.
21. Hull, C. L. *Aptitude testing*. Yonkers, New York: World Book Company, 1928, chapter 4.
22. Hull, C. L., & Montgomery, R. B. An experimental investigation of certain alleged relations between character and handwriting. *Psychol. Rev.*, 1919, 26, 63-74.
23. Hunt, T. The measurement of social intelligence. *J. appl. Psychol.*, 1928, 12, 317-334.
24. Klineberg, O., Asch, S. E., & Block, H. An experimental study of constitutional types. *Genet. Psychol. Monog.*, 1934, 16, 139-221.
25. Kretschmer, E. *Physique and character*. New York: Harcourt, Brace and Company, 1925.
26. Lewin, K. *A dynamic theory of personality*. New York: McGraw-Hill Book Company, Inc., 1936.
27. Lovell, C. A. A study of the factor structure of thirteen personality variables. *Educ. Psychol. Meas.*, 1945, 5, 335-350.
28. Moore, H. T., & Gilliland, A. R. The measurement of aggressiveness. *J. appl. Psychol.*, 1921, 5, 97-118.
29. Moreno, J. L. *Who shall survive?* Washington, D.C., *Nervous and Mental Disease Monograph Series*, No. 58, 1934.
30. Mosier, C. I. A factor analysis of certain neurotic tendencies. *Psychometrika*, 1937, 2, 263-286.
31. Murphy, L. B. *Social behavior and child personality*. New York: Columbia University Press, 1937.
32. Murray, H. A., et al. *Explorations in personality*. London: Oxford University Press, 1938.
33. North, R. D. An analysis of the personality dimensions of introversion-extroversion. *J. of Pers.*, 1949, 17, 352-367.
34. O.S.S. Assessment Staff, *Assessment of men*. New York: Rinehart & Company, 1948.
35. Paterson, D. G. *Physique and intellect*. New York: The Century Co., 1930.
36. Rorschach, H. *Psychodiagnostik*. Berlin: Hans Huber, 1932.
37. Sears, R. R. Survey of objective studies of psychoanalytic concepts, *S.S.R.C.*, 1943, 51.
38. Sheldon, W. H. and Stevens, S. S. *The varieties of temperament*. New York: Harper & Brothers, 1942.
39. Symonds, P. M. *Diagnosing personality and conduct*. New York: Appleton-Century-Crofts, Inc., 1931, chapter 5.
40. Terman, L. M., & Miles, C. C. *Sex and personality*. New York: McGraw-Hill Book Company, Inc., 1936.
41. Terman, L. M., & Oden, M. Correlates of adult achievement in the California gifted group. *Thirty-ninth Yearbook, National Society for the Study of Education*, 1940, 74-84.
42. Thurstone, L. L. A scale for measuring attitude toward the movies. *J. educ. Res.*, 1930, 22, 89-94.
43. Thurstone, L. L., & Chave, E. J. *The measurement of attitude*. Chicago: University of Chicago Press, 1929.
44. Wolff, W. *What is psychology*. New York: Grune & Stratten, 1947, chapter 13.

10

MCDUGALL AND SOCIAL PSYCHOLOGY

1. MCDUGALL'S CONTRIBUTION TO SOCIAL PSYCHOLOGY

SOCIAL PSYCHOLOGY deals in a broad sense with the behavior of individuals in groups—of the influence of the group upon the individual, and of the influence of the individual upon the group. McDougall's *Introduction to social psychology*,¹ published in 1908, was one of the first books in psychology to contain the description "social" in the title. This book was also the first to provide a systematic theoretical framework for the understanding of social behavior. William McDougall, an Englishman by birth, had done excellent work in biology, medicine, and anthropology before devoting himself to psychology. McDougall was a lecturer in psychology at Oxford University from 1904 to 1920. In 1920 he came to the United States where he was successively professor of psychology at Harvard and at Duke Universities. Besides his work in social psychology, McDougall made important contributions to abnormal psychology and to psychological theory. Strongly opposed to a "mechanistic" behaviorism, he championed instead a purposive or *hormic* (from the Greek work *hormé* meaning urge or impulse) point of view in psychology. McDougall believed that behavior is always goal-directed, variable, and persistent; that it is terminated when the goal is reached; and that purposive activity improves with repetition or practice. For a number of years McDougall was the center of a controversy between those psychologists willing to admit purpose into psychology at least with qualifications, and those who sought to reduce goal-seeking to a mechanistic stimulus-response formula.

Social behavior for McDougall was dependent upon the action of fundamental motives or drives called *instincts*. An instinct as McDougall defined it was a striving toward some goal rather than a stereotyped pattern of behavior such as is observed in animals. Also, according to McDougall,

¹ "A veritable landmark in the history of psychological development . . . it may well be asserted that the influence of that work has been felt in every branch of modern psychology." Dr. James Drever in Bernard Hollander's "McDougall's Social Psychology Anticipated by 100 Years. A Contribution to the History of Philosophy." *The Ethological Journal*, London, October 1924.

every primary instinct has a characteristic emotional core. Sometimes the motive or striving can be separated from the emotion, but in McDougall's view the two processes are ordinarily so interwoven as to be hardly distinguishable.

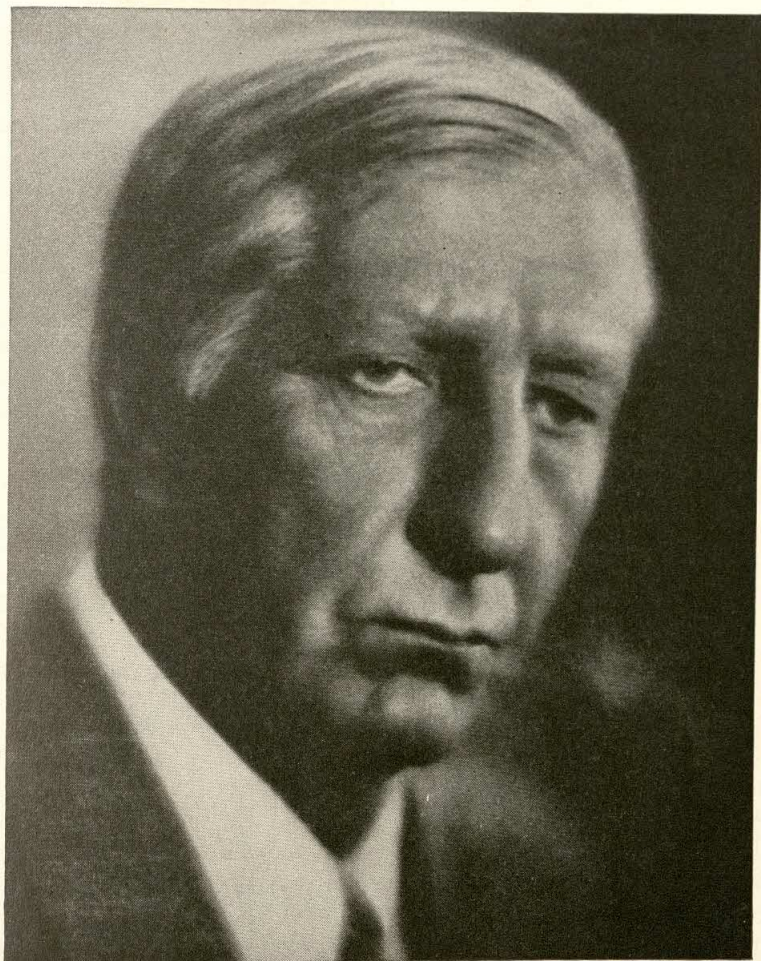
McDougall's instincts or native propensities provided the ground work for his social psychology. In 1908 in his "Introduction" he listed twelve major human instincts and their accompanying emotions; by 1932 the list had been revised to include seventeen major instincts beside a number of other activities such as breathing, sneezing, crying, sleeping, and the like. McDougall's main list of instincts or propensities and their characteristic emotions is given below. He believed it to be instinctive:

INSTINCTS	ACCOMPANYING EMOTIONS
1. To desire food	Appetite or craving
2. To reject certain substances or things	Disgust or repugnance
3. To explore new places and things	Curiosity
4. To escape from danger	Fear
5. To fight when challenged	Anger, rage
6. To have sex desire	Lust or sex excitement
7. To care for the young (mothering)	Tender emotion, love
8. To seek company (gregariousness)	Loneliness, nostalgia
9. To seek to dominate (self-assertion)	Elation, positive self-feeling
10. To accept inferior status (submission)	Humility, negative self-feeling
11. To make things	Feeling of creativity
12. To collect things	Acquisitiveness, pride of possession

McDougall wrote regarding his list of instincts as follows (Introduction, 1908, p. 44): "Directly or indirectly the instincts are the prime movers of all human activity . . . determine the ends of all activities and supply the driving power . . ." While fundamentally these instincts were thought to be innate, they become modified by experience and learning and their expression depends upon the character of the environment. In adulthood McDougall believed the instinctive tendencies were modified, organized, and often combined into *sentiments*. Religion, patriotism, friendships, loyalties to school or community, and other complexes of emotions, values, and attitudes fall under the head of sentiments.

2. INSTINCTS IN SOCIAL PSYCHOLOGY

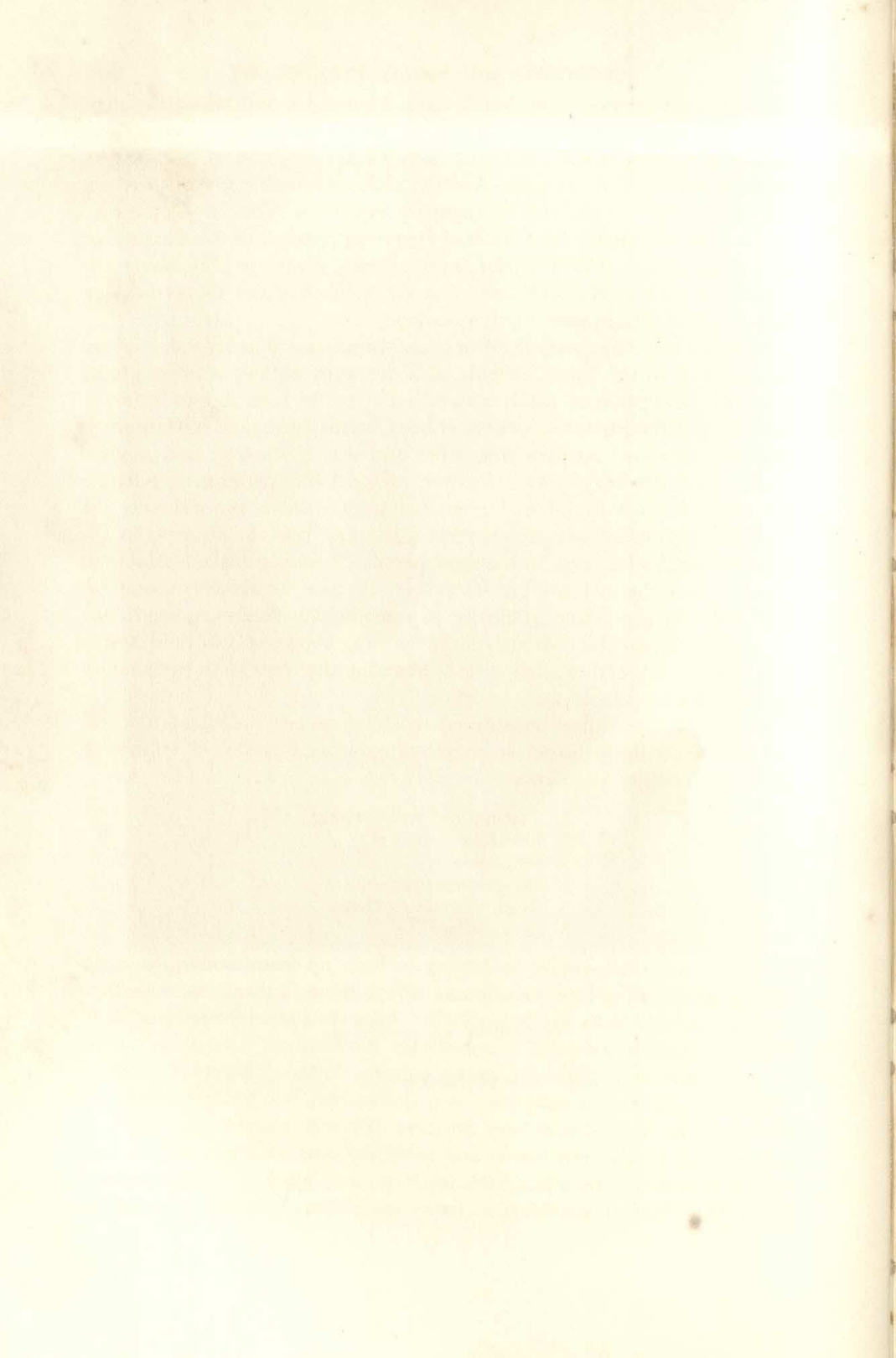
McDougall's view of social behavior as determined by innate drives or instincts soon ran into difficulties and is no longer widely accepted by social scientists. Criticism of the explanatory value of the instinct hypothesis was expressed in 1924 by Bernard, who showed that the lists of in-



Press Association, Inc.

WILLIAM MCDOUGALL

1871-1938



instincts drawn up by James, McDougall, Thorndike, and others included almost every activity from sneezing to "love of the sea." Nearly 6,000 human activities, in fact—some obviously learned—have at one time or another been called instincts. Another difficulty with the term instinct is its circularity: it explains behavior by naming it. Thus, if people congregate it is because of the instinct of gregariousness; while it is the instinct of gregariousness which causes them to seek company. The same circularity is seen in the statement that children construct things because they have an inborn constructive tendency.

Behavior which repeats itself in a consistent and stereotyped way can be observed in the lower animals, and the term instinct is undoubtedly a useful description of much animal behavior. In fact, it was from observations of the ritualistic activity of bees, wasps, birds, and other animals that the notion of instincts first arose and was taken over and applied to man. But stereotyped and clearly distinguishable patterns of behavior are rarely if ever found in human beings. Consider, for instance, the so-called instinct of self-preservation ("the first law of nature"). In the face of imminent danger to life most people do indeed make a strenuous effort to save themselves—but *what* they do does not always accomplish this end. On the whole, difficulty in defining an instinct in man, lack of stereotypy and the obviously large learning component in most activities has forced social psychologists to abandon the term as an explanatory concept in group behavior.

In 1942 F. H. Allport substituted what he called "prepotent reflexes" for instincts. Allport listed six important groups of prepotent reflexes in the human being, as follows:

1. Starting and withdrawing
2. Rejecting
3. Struggling
4. Hunger reactions
5. Sensitive zone reactions
6. Sex reactions

Social behavior, according to Allport, is built up from modifications on the stimulus side and on the response side of these fundamental activities. While less rigid than McDougall's list, Allport's prepotent reflexes have not been widely accepted as substitutes for instincts. Human behavior is modifiable to a high degree by culture. Anthropological studies of primitive peoples, for example, have shown that two groups which are racially alike may exhibit very different attitudes toward child rearing, say, as well as different family and tribal customs. If behavior were determined primarily by a few basic instincts, such marked variations could hardly arise. Social psychologists today are inclined, therefore, to ascribe

group behavior, whether in the family or in a nation, to motives or drives which are in large part learned or culturally determined.

3. FACTORS IN SOCIAL BEHAVIOR

Various factors have proved to be important determinants of social behavior. Several of these have been studied experimentally and brief accounts of this work will be given in the present section.

(1) The effects of *prestige suggestion* in determining what one perceives have been studied in a series of experiments which made use of the "autokinetic" effect. Autokinetic effects, in part physiological and retinal phenomena, can be readily obtained in the following way. Let a person who is in complete darkness observe a single pin point of light, and the light will appear to move or drift in somewhat erratic fashion. A single point of light cannot be localized when seen in the dark because there are no objects in terms of which to anchor or fix it. The distance through which the light moves will vary considerably from person to person. Thus Sherif found that a subject will establish a range or norm for the movement he perceives, e.g., 2 inches or 6 inches, say; and that this figure remains fairly constant in subsequent trials. However, when two or more individuals who had established different norms when alone were brought together, there was a strong tendency for a new or compromise norm to be set up. Furthermore, a common group norm was set up when the apparently moving light was observed by individuals working together for the first time. From several experiments it seemed clear that the shift in a subject's "frame of reference" or standards of judgment toward conformity with the group could be explained by prestige suggestion. Many observers were not aware that their judgments had been influenced by the presence or the remarks of the experimenter or by a member or members of the group. Others were aware of the suggestion effect. Interestingly enough, an observer often continued to show conformity to the social norm previously set up, even when his judgments were made later on in the alone condition.

(2) *Aspiration-level* refers to the goal of achievement or performance which a person sets up for himself and hopes to realize. An individual's aspiration-level will depend not only upon his own standards but also upon those developed by the group of which he is a part. In one experiment designed to test the effects of suggestion upon aspiration-level, four groups, A, B, C and D, were set up with about 20 in each group. All subjects were told that they would be given a test of fifty questions arranged in multiple-choice form and dealing with knowledge of literature. The following item served as a sample:

David Copperfield was written by (1) Thackeray
 (2) Dickens
 (3) Thomas Hardy

Group A received no special instructions; Group B was told that authors and literary critics had made an average score of 37.2 on the test; Group C that other students in psychology had made an average score of 37.2; and Group D that a number of unemployed workers had made an average of 37.2. In addition, each subject was told that the maximum score was 50 and that by guessing (since there were three choices) one should get a score of about 17.

Before taking the test the members of each group were asked to indicate what scores they expected to make. It may be repeated that Group A had no instructions, B was told what score the experts had made, C what score a similar group had made, and Group D what presumably inferior (unemployed) workers had done. The mean aspiration-levels of the four groups were as follows:

	Groups			
	A	B	C	D
Mean aspiration-level	26.95	23.09	31.09	33.05

These results show clearly the effect of suggestion upon aspiration-level. Group B, challenged by the performance of a supposedly superior group, set its aspiration-level low; while Group D set its aspiration fairly high, confident of its ability to equal (almost) the performance of an inferior group. Group A serves as a sort of low reference point and Group C set its mean aspiration-level fairly close to that of a supposedly similar group. It has been suggested that the same aspiration-level may arise from quite different factors in different individuals and that knowledge of the individual is necessary in order to interpret the goals which he sets. Furthermore, it seems clear that the aspiration-level of a group can be altered somewhat more readily than the aspiration-level of an individual, perhaps because of the division of responsibility within the group.

(3) *Social facilitation* is said to exist when an individual's performance is affected favorably by the presence of a group of which he is a part. The facilitating effect of the group has been frequently reported. In one experiment, for example, subjects gave more word associations when working together than when working alone. In another series of experiments, social facilitation as well as prestige suggestion apparently operated. Subjects were asked to report upon the pleasantness and unpleasantness of various odors when reports were given singly and in groups. There were fewer extreme judgments when the individual worked in the group, the presence of other observers making judgments easier as well as bringing

them closer together. In still another experiment, an attempt was made to separate the factor of *competition* from that of social facilitation. Subjects were given the task of multiplying two-place by two-place numbers, completing analogies, and providing word associations in serial order. The experimenter wished to make comparisons between an individual's achievement (1) when working alone, (2) when working with others but non-competitively, (3) when striving to excel others in the group as rivals, and (4) when working under the close observation of spectators. While the presence of an audience had a decidedly facilitating effect upon the speed of work though not upon its accuracy, working alone with a non-competitive group had little if any effect upon the individual's performance. Increases in score due to rivalry or competition were greater than the increases in performance due to social facilitation alone indicating that competition is a strong incentive to improved performance in most social situations.

Several of the early social scientists, greatly impressed by the distinctive character of group behavior, coined the expression "group mind" or "crowd mind." This term implied the existence of an overall will or directing mind which is not simply the sum of the individual minds in the group. The concept of a group mind is an interesting speculation, but it is vague and somewhat mystical and is not generally accepted today. While admitting that group behavior has its own peculiar patterns and is not simply a sum-total of individual behaviors, present-day sociologists and social psychologists see no reason to posit the existence of an emergent "group mind." Rather they study the interactions of individual minds in their creation of group behavior.

(4) Various social scientists have stressed the importance of *rôle* and of *status* in social behavior. The place in a given social scheme which an individual occupies serves to define his social status or position. *Rôle*, as distinct from status, includes the behavior patterns which identify and are associated with a particular status. A physician, a clerk in a shoe store, and a street sweeper—each has a definite status within the social system of his community. In addition, each of these people has a *rôle* which depends upon his status. Illustrative of well-known rôles are the "bedside manner" of the physician, and the professional "friendliness" of the politician. A businessman of unassuming manner and modest tastes, if elected mayor of his town or state senator, say, will often assume the "character" or *rôle* which he thinks is appropriate to his new status. Sociological studies of American community life have shown the effects upon behavior of economic as well as social status. Within most American cities, membership in the upper, middle, or lower classes depends in varying degree upon money, upon "family" and education, and upon personal initiative

and enterprise. From one social class to another, standards of good manners and ethical practices differ as well as attitudes and beliefs toward social and economic issues. Among other things, for example, middle class children have been reported to be more pampered and less able to assume responsibility than children in lower class families.

Status within a school class is sharp and clearly defined. In studying preferences among classmates, for example, Moreno (p. 211) found that certain children were selected as desirable seat mates or companions far more often than others. Athletic prowess, leadership, physical size, and personal qualities play a part in these choices.

(5) The tendency of an individual to accept the customs of his community, that is, to show *conforming behavior*, has been of interest to many social psychologists. Conformity by an individual may result from prestige suggestion—as when a person patterns his behavior after those whom he respects (p. 203). Again, an individual may be unfamiliar with other practices, or he may not wish to be regarded as queer or different, or lack of conformity may lead to punishment or social disapproval. F. H. Allport has shown that conforming behavior often follows a curve which may be compared roughly with the reversed letter *J*, but without the tail of the *J* being turned up. Figure 35 shows what was done by 2,114 motorists whose characteristic behavior was observed at a boulevard stop sign marking an intersection. It is obvious that the majority (1,544) stop; but a fair number (462) proceed very slowly, while a few (58) slow down slightly or not at all. Conformity behavior in this situation is well represented by a *J*-curve. When there is a red light at an intersection as well as a traffic policeman, conformity is considerably higher and the *J*-curve steeper. Conforming behavior has been studied in a variety of situations: parking in time-limited zones, punching the time-clock in a factory, attendance upon lectures in college. In all of these cases there is a piling up at the one end of the distribution (the conformity end) with a decreasing but often considerable number showing non-conforming behavior. Conforming behavior is apparent in various crowd phenomena—in the razzing of the umpire at a ball game, for instance, or in the noisy adulation of 'teenagers for a currently popular actor or singer.

(6) The *frustration-aggression* hypothesis expresses the fact that aggressive behavior may arise from a prior frustration or curtailment of an individual's activities. Evidence for this idea rests upon observation, clinical case histories, and a few experimental studies. In one experiment 6 students were deprived of sleep for twenty-four hours during which they were annoyed and badgered in various ways: prohibited from smoking, deprived of food, promised and then denied a chance to play cards, promised and refused a hot breakfast, etc. Definite hostility toward the

experimenters resulted from these "frustrations" but the only direct aggression was in the form of sarcastic references, glumness, and some rudeness. In another well-known experiment, children were first allowed to play with attractive toys and were then frustrated by being allowed to view the toys but not to use them. A frustrating situation of this kind reduced the constructiveness of play and interfered markedly with the child's spontaneity of behavior.

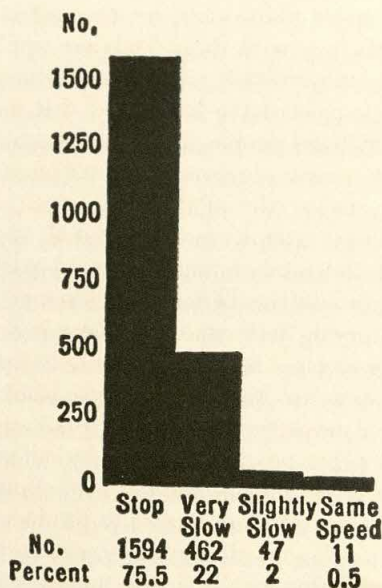


FIG. 35

DICKENS MOTORISTS' BEHAVIOR AT STOP SIGN AND CROSS TRAFFIC CORNERS
(After Newcomb, Hartley, and Others, *Readings in Social Psychology*, 1947. Henry Holt and Company.)

Apparently, children can be taught to react to frustration in other than an aggressive manner. In one study, for example, 82 preschool children 3 to 6 years old were given the task of placing blocks in a puzzle box and were observed for reaction to failure. While interesting, the task was difficult to complete in the fifteen minutes allowed. Some of the children gave up, some asked for help, and some showed marked emotional reactions to failure—by crying, angry expressions, sulking. Training was now administered to 12 of the 15 children who had shown immature emotional reactions to failure in the puzzle situation. Training consisted of giving each child a series of progressively more difficult problems, thus leading

him to the solution of a difficult problem by easy steps. Upon retesting most of the immature behavior (sulking and crying) previously shown dropped out. After training, too, the children tried harder, showed more interest in solving the problems themselves, and exhibited a minimum of exaggerated emotional behavior.

4. METHODS IN SOCIAL PSYCHOLOGY

Groups are made up of individuals, hence most of the methods used in experimental psychology with individuals are applicable to the study of groups. Systematic *observation* is a standard method in social psychology, and anthropologists have long used this method in studies of primitive and semi-civilized people. For example, considerable variation in child-rearing practices was observed by Mead among neighboring New Guinea tribes that were racially of the same stock. Marked personality differences among three groups were attributed in large part to early training, the rôle of the mother being deemed particularly important. Psychologists have studied the personal and social growth of children by systematically observing and measuring them in a variety of situations over a number of age levels. The case history falls also under the method of systematic observation. In constructing a case history, quantitative data (test scores and other measures), as well as qualitative estimates of personality and character, are nearly always included.

The purpose of the *experimental method* is to allow us to evaluate and measure the power of one or more factors to produce significant changes in behavior. In the psychological laboratory, extraneous variables which interfere with the operation of the factors under study are allowed for or are controlled. It is difficult, however, to isolate the crucial variables operative in many areas of social psychology; and in such problems as the causes of juvenile delinquency or the reasons for riots or panics, it is impossible to exercise experimental control. In spite of such limitations we have seen (pp. 202-206) that at least some of the determinants of social behavior have been studied under experimental conditions; and other systematic experimental approaches to social phenomena will be found in Section 5.

Opinion-polling is a technique peculiarly well suited to the study of social behavior. In its simplest form a poll consists of one or more questions put to a group of people. Great care must be exercised in order to get a sample which will be a representative cross section of the larger group in which we are interested. Thus to secure a nation-wide sample, selection must be made in terms of age, sex, economic status, population

size of state or community, and rural or urban residence. A representative group must not contain too many old people or too few men or too many from populous states or large cities.

Size of sample is an important consideration, though contrary to common opinion sample size is not as significant a factor as representativeness of sample. The classical example of a forecast which went wrong though based upon a large vote is the *Literary Digest's* prediction of Roosevelt's defeat by Landon in 1936. The *Literary Digest's* poll was based upon approximately two million ballots. But despite its large size, this sample was biased, as the *Digest* relied heavily upon lists of telephone subscribers and automobile owners—people whose names are easy to get but who fall in the upper income brackets. The failure of the polls to predict the election of Truman in 1948 has been explained by a variety of factors: closeness of the election, errors of sampling in some states, failure to allow for the undecided voters, and failure to measure preferences just before the election. One factor, the closeness of the election, was probably crucial. It has been pointed out, for instance, that Dewey could have won by carrying Ohio, California and Illinois, in which three states he lost by less than 1 per cent of the vote.

It is possible from probability tables to determine the size of sample necessary to secure a specified level of accuracy. For example, if opinion in the population divides about equally (50-50), the mathematical chances are 99 in 100 that a random sample of 900 cases will give an accurate prediction within a range of 5 per cent. If opinion divides in the ratio of 80 to 20, only 576 cases are needed to obtain a forecast accurate within 5 per cent.

It is important to plan the questions so that the respondent will know exactly what is wanted. Involved questions must be avoided as well as those which are ambiguous or "loaded" in a given way through the use of prestige and suggestion factors or emotional terms like Red or Fascist. Usually it is advisable to pretest the questions, that is, try them out on small samples. Obscurities and misunderstandings can be cleared up many times only through a preliminary trial. One of the best-known pollsters (Gallup) has suggested the following criteria for questions:

1. Questions should be brief and to the point. Omit conditional and dependent clauses.
2. Use simple words and phrases in everyday use. For example, instead of asking people whether nations should arbitrate their difficulties, say "settle by conference."
3. Omit words which have strong emotional bias.
4. If choices are lengthy or numerous, print them on a card which the respondent can read.

In general it is better to use questions which are positive instead of negative, objective instead of subjective. To illustrate, it is better to ask "Is it desirable to permit installment buying" than to ask "Is it desirable to prohibit installment buying"; to ask "Is it desirable" than to say "Do you feel that it is desirable."

5. SYSTEMATIC STUDIES OF SOCIAL BEHAVIOR

A number of investigations have been concerned with the identification and measurement of those variables or factors which condition social behavior. While some of the results are necessarily qualitative, many are quantitative and depend directly upon the use of such psychological instruments as rating scales, questionnaires and psychological tests. The following are samples of the experimental work in this field.

(1) *Attitudes and Stereotypes*. Attitudes may be described as general tendencies to behave in certain consistent ways toward people, institutions, happenings. Attitudes often become stereotyped—that is, fairly rigid and standardized. Most of us, for example, are inclined to catalogue the people whom we meet as intellectuals, foreigners, lowbrows, radicals. Stereotypes provide an easy way of making judgments, and as a result are relinquished with great reluctance.

In his studies of "social distance" Bogardus was concerned indirectly with stereotypes. Bogardus asked a group of 1,725 Americans to check the following propositions with regard to forty different racial or national groups:

1. Would admit to close kinship by marriage
2. Would admit to my club as personal chum
3. Would admit to my street as neighbor
4. Would admit to employment in my occupation
5. Would admit to citizenship in my country
6. Would admit as visitors only to my country
7. Would exclude from my country

Note that the social distance or degree of acceptance decreases steadily from 1 to 7. Table XXI shows results for five of the highest ranking and five of the lowest ranking groups.

While 93.7 per cent of the respondents indicated a willingness to marry people of English stock, only 1.4 per cent were willing to marry Negroes. In general, North Europeans rank higher than South Europeans, probably because the respondents are mostly of North European ancestry.

Other studies (Katz and Braly) have shown that social distance depends to a high degree upon the stereotypes which we attach to racial and national groups. College students, for example, most often described

Germans as "scientifically minded" and "industrious," English as "sportsmanlike" and "intelligent," Jews as "shrewd" and "mercenary," Negroes as "superstitious" and "lazy" and Turks as "cruel" and "very religious."

TABLE XXI
PERCENTAGES OF 1,725 AMERICANS CHECKING EACH OF PROPOSITIONS
GIVEN ABOVE (Bogardus)

	Propositions						
	1	2	3	4	5	6	7
English	93.7	96.7	97.3	95.4	95.9	1.7	.0
Canadians	86.9	93.4	96.1	95.6	96.1	1.7	.3
Scotch	78.1	89.1	91.3	92.8	93.3	1.7	.0
Irish	70.0	83.4	86.1	89.8	91.4	4.0	.7
French	67.8	85.4	88.1	90.4	92.7	3.8	.8
Negroes	1.4	9.1	11.8	38.7	57.3	17.6	12.7
Turks	1.4	10.0	11.7	19.0	25.3	41.8	23.4
Chinese	1.1	11.8	15.9	27.0	27.3	45.2	22.4
Korean	1.1	10.8	11.8	20.1	27.5	34.3	13.8
Hindus	1.1	6.8	13.0	21.4	23.7	47.1	19.1

Attitudes within social groups have been studied by Moreno's "sociometric" technique. The social structure in a first grade, for example, is represented graphically in Figure 36. Note that of the 21 boys and 14 girls who were asked to choose from among their classmates those whom each wanted to have stay in the same classroom and sit near him, 18 are unchosen, 5 are "stars" or centers of attraction, and 6 form 3 pairs. The number of individuals isolated, that is, unchosen by their classmates, was highest in the kindergarten and decreased gradually to the seventh grade. An interesting finding is that the number of boys choosing girls and girls choosing boys is highest in the kindergarten and first grade and lowest in the fourth and fifth grades. After the fifth grade, choices between the sexes gradually increase.

The sociometric technique has been used by Jennings to study the social attitudes of a group of girl adolescents committed by the Children's Courts to the New York State Training School for Girls. These girls were of normal intelligence, fell between the ages of 12 and 16, and represented a cross section of the socially and economically underprivileged in the state. Each of the girls in the experimental sample was asked to designate those companions with whom she preferred to live, to work, to study, and to spend her recreation time. The girls were also asked to designate girls whom they should not like to select for these activities. Girls chosen more often than average were called "over-chosen"; those chosen less often than average "under-chosen." The over-chosen girls were

found to be those who took the initiative or assumed leadership more often than the average and who frequently showed ingenuity in fostering friendships among new members. They were inclined also to be more

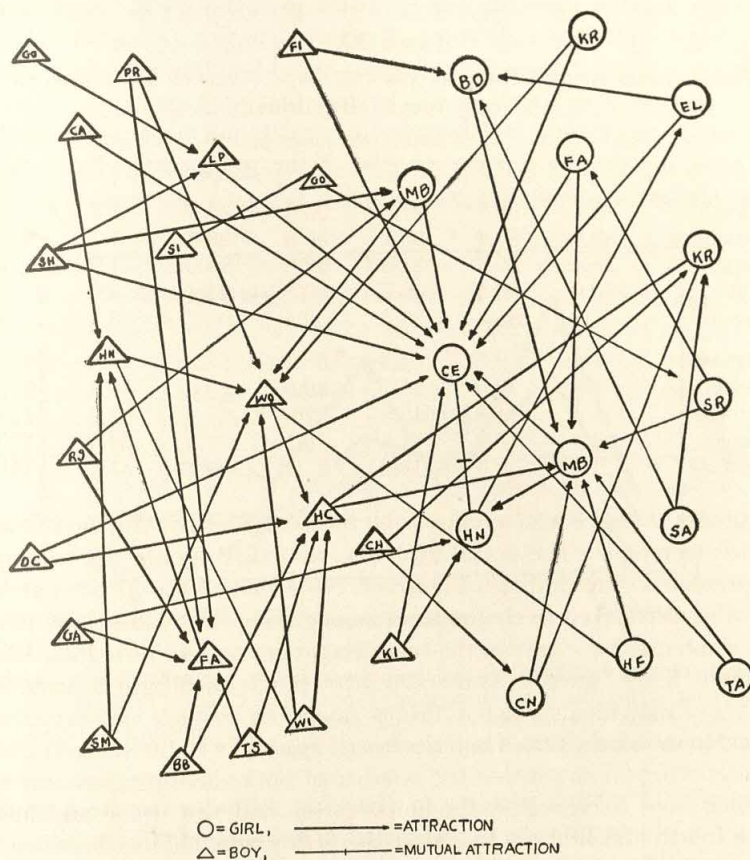


FIG. 36

CLASS STRUCTURE, FIRST GRADE

Class of 21 boys and 14 girls. Unchosen, 18; pairs, 3; stars, 5; chains, 0; triangles, 0; intersexual attractions, 22. (After Moreno, *Who Shall Survive?* 1934.)

rebellious than the average, perhaps because of their greater enterprise and vigor. Under-chosen girls ranked much lower in all of these respects than the over-chosen.

(2) *Propaganda and the Movies.* Any deliberate attempt to change a person's opinions, attitudes or beliefs and thus influence his behavior can perhaps be described as propaganda. Within this meaning of the term, advertising and even education are propaganda agencies as they are concerned primarily with changing behavior. Propaganda techniques

include face-to-face argument, books, pamphlets, lectures, radio, motion pictures, and television.

Unfortunately a great many people think of propaganda as something inherently bad or immoral, just as habit to many means "bad habit." According to this view only those efforts at changing a person's behavior which are vicious, false and misleading should rightly be called propaganda. In answer to this it should be said that propaganda in the sense of attempts at attitude and behavior control is not *necessarily* good or bad. It is, in fact, the purpose or goal of the propaganda which determines whether its effects are good or evil.

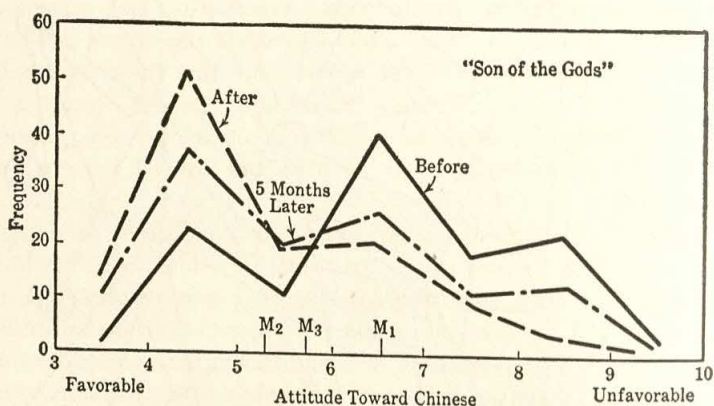


FIG. 37

Attitude of school children toward the Chinese before, immediately after, and five months after seeing the pro-Chinese film, "Son of the Gods." (From Peterson and Thurstone, *Motion Pictures and the Social Attitudes of Children*, The Payne Fund Motion Picture Studies, 1933. The Macmillan Company.)

The rôle of motion pictures in changing attitudes has been studied in a series of experiments. In one instance, five motion picture films were used including one favorable to and one unfavorable to the Chinese, a pro-German film, an anti-gambling film and one designed to show the evils of bootlegging. Subjects were high school children in small towns in Illinois. The method used was to administer an attitude questionnaire before and after the picture. In the picture favorable to the Chinese (called "Son of the Gods") the attitude of the children toward the Chinese was remeasured the day after the picture was shown. In one group the attitude scale was administered after five months and in still another group after nineteen months. Figure 37 shows a decided shift toward a more favorable attitude as a result of the picture. A good deal of this shift persisted after five months; and more than half of the change remained after a year and a half.

(3) *The O.S.S. Project "The Assessment of Men."* Undertaken by the Office of Strategic Services during World War II, this project represents a comprehensive and at the same time intensive attempt to evaluate the capabilities of a group of men for special and secret services. The task given the Assessment Staff of O.S.S. was that of estimating suitability for a variety of tasks. Some of these were administrative, others required men who could be leaders of guerrilla units, saboteurs, propagandists, radio speakers, "idea men," and the like. Most assignments demanded a high degree of intelligence, physical stamina, courage, loyalty, and devotion to duty. Over a period of two years the Assessment Staff of some seventy-five psychologists, psychiatrists, sociologists, and other experts studied and evaluated more than 5,000 persons at the rate of 200 or more a month. While the authors of the report state that their approach was "organismic," this means little more than that an earnest effort was made to get as complete a picture as possible of each assessed person—not only particular aptitudes and special gifts but overall personality and character traits.

For a period of four days, small groups of candidates lived together with the psychologists and psychiatrists on an estate near Washington, D.C. During this time the candidates were carefully observed, rated, interviewed, given questionnaires and put through various aptitude and "situation" tests. A wide variety of tests and techniques were drawn upon: intelligence tests (Wechsler-Bellevue), projective tests (Rorschach, TAT), questionnaires, attitude scales, personal history forms. Ratings were used to evaluate motivation, drive and personality characteristics and also to aid in informal observations of behavior. Finally, a number of situational tests posed problems intended to reveal a candidate's imagination, initiative, coöperation, leadership, cool-headedness, and ability to surmount difficulties.

The situation tests were often ingenious. In the "brook task," for example, a group of four to seven candidates at a time were told to imagine that an ordinary brook was a raging torrent over which a delicate piece of apparatus (really a log) and a box of ammunition (100-pound rock) must be transported. The problem was to work out ways and means of doing this with available equipment—several boards all too short, pieces of rope, a pulley and an empty barrel without ends. Candidates were rated for coöperation, leadership and initiative as shown in the solution of this task. Other situations required the crossing of barriers, the directing of construction tasks (to show leadership) and the undertaking of various physical tasks which might indicate courage, stamina, and self-control. A "stress interview" subjected the man to an intensive cross examination, the purpose of which was to see how well he could defend

himself and stick to a story (which he had previously composed) under a barrage of questions.

Records on each man consisted of scores on the tests, assessments for such personality traits as motivation, practical intelligence, emotional stability, social adjustment and stamina, personal and medical history, and an overall interview. On the basis of all of these data a concise summary of the individual's assets and liabilities was constructed.

While it was impossible to make a full appraisal of the validity of many assessments, the authors of this study do report some data upon the validity of their procedures. In view of the thoroughness with which these assessments were made, their validity is probably much higher than often appeared on the surface.

(4) *Attitude Measurement Through Opinion-Polling Techniques.* Studies by the U.S. Strategic Survey (Morale Division) of the changes which took place in Japanese civilian morale as result of bombing raids offer a good illustration of the usefulness of opinion-polling. A set of about 40 questions was drawn up covering the probable reaction of Japanese civilians to the effects of bombing and to other events of the war. The work of interviewing was completed in 1945, four months after the surrender. Stratified sampling was followed so that the interviewees constituted a representative cross section of the Japanese population for age, sex, and education. In addition to formal answers to the questions, each interviewer took full notes which were later written up in the form of a report. Figure 38 shows the rapid growth of unwillingness on the part of most Japanese to continue the war. Loss of morale was rapid after low-level air raids on major cities and by June 1945 more than 60 per cent of the population were unwilling to continue the war.

The technique of opinion-polling was also used during World War II to study the attitudes of the American soldier toward the Army, his training, discipline, officers, knowledge of circumstances leading up to war, and the goals for which his country was fighting. These data are valuable to the military authorities as they provide a factual basis for administrative policies. Also they are a gold mine of information to the social scientist.

During World War I the policy was followed of placing Negro and white soldiers in separate units, and this plan was continued with some modifications at the beginning of World War II. In several opinion polls interesting data were obtained upon the attitudes of white and Negro soldiers toward service in the same outfit. In 1943 the question "Do you think white and Negro soldiers should be in separate outfits or should they be together in the same outfit?" was proposed to 4,800 white and 3,000 Negro enlisted men. These groups while small were believed to provide a cross section of opinion in the two races. Results shown in

Table XXII indicate that among the white soldiers the majority opinion was in favor of separate outfits for Negro and white. Negro soldiers, on the other hand, were about equally divided in their attitudes toward separate and integrated services.

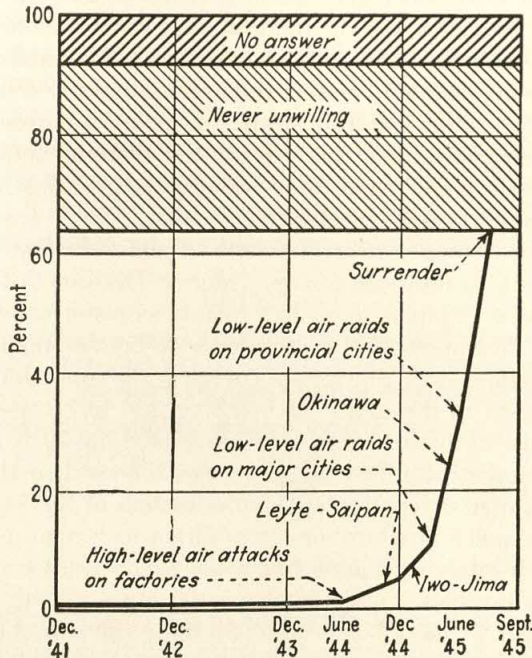


FIG. 38

GROWTH OF UNWILLINGNESS TO CONTINUE THE WAR

This chart shows the percentages of Japanese who said they were unwilling to continue fighting, at various stages of the war. The chart is based on answers to the following questions: "Did you at any time during the war come to a point where you felt you could not go on with the war? When was that?" (This is No. 18 of the question list, followed by the request to specify the time to which the answer referred. (From Crafts, Schneirla, Robinson, and Gilbert, *Recent Experiments in Psychology*, 2nd ed. Copyright, 1950. McGraw-Hill Book Company, Inc.)

TABLE XXII

ANSWERS IN TERMS OF PERCENTAGES GIVEN BY 4,800 WHITE SOLDIERS AND 3,000 NEGRO SOLDIERS TO THE FOLLOWING QUESTION: "DO YOU THINK WHITE AND NEGRO SOLDIERS SHOULD BE IN SEPARATE OUTFITS OR SHOULD THEY BE TOGETHER IN THE SAME OUTFIT?"

	White	Negro
Should be separate	84%	37%
Makes no difference	4	17
Should be together	12	36
Undecided	0	10

Toward the end of the war a number of platoons made up of Negro volunteers were attached to white companies and served with them in combat. The favorable opinion of white officers and men toward Negroes after experience with them as combat soldiers is shown clearly in Table XXIII. On the whole the Negro platoons attached to white companies were younger, somewhat more intelligent, and better educated than the rank and file of Negro troops, but the differences were not great. Favorable attitude after combat experience toward the attachment of Negro platoons to white companies is also shown by the fact that in 1945, 69 per cent of the officers and 83 per cent of the sergeants and men said the Negro troops were equal to white soldiers of the same training and experience. Furthermore, 55 per cent of the officers and 72 per cent of the enlisted men thought it was a good idea to have Negroes in the Infantry.

TABLE XXIII

ANSWERS IN TERMS OF PERCENTAGES OF WHITE COMPANY OFFICERS AND WHITE PLATOON SERGEANTS AND ENLISTED MEN TO THE QUESTION, "HOW WELL DO COLORED SOLDIERS IN THIS COMPANY PERFORM IN COMBAT?"

	White Company Officers	White Platoon Sergeants and Enlisted Men
Very well	84%	81%
Fairly well	16	17
Not so well		1
Undecided		1

It seems evident that the stereotype of the Negro (p. 209) as happy, dull and shiftless breaks down when white and Negro soldiers fight side by side. The increased respect which white soldiers have for Negroes upon closer acquaintance is shown by the fact that 62 per cent of white soldiers in companies in which there were *no* Negroes disliked the idea of mixing Negro and white troops, whereas only 7 per cent of white soldiers in companies with Negro platoons disliked the idea of mixed groups.

(5) *Social Factors and the Productivity of Workers.* A series of studies carried on in the Western Electric Company have shown clearly the importance of social and psychological factors in worker satisfaction and productivity. The object of the studies was to get information on the effects of differences in illumination and temperature upon output as well as the influence of rest periods, wage rates and other factors. The expectation was that variations in output would be associated with variations in work conditions. A group of six girls whose task was the assembly of telephone relays was selected for the experiment. These girls were described as average workers, neither inexperienced nor expert; their

work was of a sort that slight changes in rate of output would be quickly revealed by the work record.

A general and unexpected result was the finding that output increased regularly whatever the experimental conditions—whether the rest periods, for example, were numerous or few, long or short. Reasons for this unexpected result seem to depend upon various social and psychological factors. For one thing, the girls felt honored to be chosen for the experiment and enjoyed working in the test room. They reported that supervision was less exact, and that they were allowed more liberties (conversation, for example). As a result a team spirit was developed, and a degree of coöperation not present before.

By contrast, in the regular or non-experimental departments the investigators found considerable discontent with wages, supervision and promotion practices together with fear of speed-up. The common feeling was that a worker should not turn out too much or too little work. If he did too much he was accused of being a "rate buster" or "speed king." If he turned out too little he was called a chiseler, while if he reported to management anything which might injure a fellow worker he was a "squealer." The conclusion of this study was that worker output and satisfaction depend to an equal or greater degree upon social and psychological satisfactions than upon wages and working conditions.

BIBLIOGRAPHY

1. Allport, F. H. The influence of the group upon association and thought. *J. exp. Psychol.*, 1920, 3, 159-182.
2. Allport, F. H. *Social psychology*. Boston: Houghton Mifflin Company, 1924.
3. Allport, F. H. The J-curve hypothesis of conforming behavior. *J. soc. Psychol.*, 1934, 5, 141-183.
4. Barker, R. G., Dembo, T., and Lewin, K. Frustration and regression: an experiment with young children. *University of Iowa Studies in Child Welfare*, 1941, 18, 1.
5. Bernard, L. L. *Instinct, a study of social psychology*. New York: Henry Holt and Company, 1924.
6. Bogardus, E. S. *The measurement of social distance, from Immigration and Race Attitudes*. Boston: D. C. Heath and Company, 1928.
7. Chapman, D. W., & Volkman, J. A social determinant of the level of aspiration. *J. abnormal soc. Psychol.*, 1939, 34, 225-238.
8. Dashiell, J. F. An experimental analysis of some group effects, *J. abnormal soc. Psychol.*, 1930, 25, 190-199.
9. Dollard, J., et al. *Frustration and aggression*. New Haven: Yale University Press, 1939.
10. Doob, L. W. *Public opinion and propaganda*. New York: Henry Holt and Company, 1948.
11. Gallup, G. H., & Rae, S. F. *The pulse of democracy. The public opinion poll and how it works*. New York: Simon & Schuster, Inc., 1940.

12. Jennings, H. H. Leadership and sociometric choice. *Sociometry*, 1947, 10, 32-49.
13. Jones, H. E. The California adolescent growth study. *J. educ. Res.*, 1936, 31, 561-567.
14. Katz, D., & Braly, K. W. Racial stereotypes of 100 college students. *J. abnormal soc. Psychol.*, 1935, 30, 175-193.
15. Keister, M. E., & Updegraff, R. A study of children's reactions to failure and an experimental attempt to modify them. *Child Development*, 1937, 8, 241-248.
16. Klineberg, O. *Social psychology*. New York: Henry Holt and Company, 1940.
17. Linton, R. *The cultural background of personality*. New York: Appleton-Century-Crofts, Inc., 1945.
18. McDougall, W. *Introduction to social psychology*, 5th ed. London: Methuen & Company, 1912.
19. McDougall, W. *The energies of men, a study of the fundamentals of dynamic psychology*. London: Methuen & Company, 1932.
20. Mead, M. *Growing up in New Guinea*. New York: William Morrow & Company, 1930.
21. Moreno, J. L. *Who shall survive? A new approach to the problem of human inter-relations*. Washington: Nervous and Mental Diseases Monog. Series, 1934, 58.
22. O.S.S. Assessment Staff. *Assessment of men*. New York: Rinehart & Company, 1948.
23. Peterson, R. C., & Thurstone, L. L. *The effect of motion pictures on the social attitudes of high school children*. New York: The Macmillan Company, 1933.
24. Robinson, A. L. *William McDougall, a bibliography*. Durham, N.C.: Duke University Press, 1943.
25. Roethlisberger, F. J., & Dickson, W. J. *Management and the worker*. Cambridge, Mass.: Harvard University Press, 1939.
26. Sargent, S. S. *Social psychology*. New York: The Ronald Press Company, 1950.
27. Sherif, M. *The psychology of social norms*. New York: Harper & Brothers, 1936.
28. Stouffer, S. A., et al. *The American soldier: adjustment during army life*. Princeton, N.J.: Princeton University Press, 1949, Vol. 1.
29. U.S. Strategic Survey (Morale Division). *Effects of strategic bombing on Japanese morale*. Washington, D.C.: U.S. Government Printing Office, 1947.
30. Woodworth, R. S. *Contemporary schools of psychology*, rev. ed. New York: The Ronald Press Company, 1948.

BINET AND THE INDIVIDUAL INTELLIGENCE TEST

1. BEGINNING OF THE MOVEMENT FOR MEASURING INTELLIGENCE

IN 1904 the French Minister of Public Instruction appointed a commission of physicians, educators, and scientists to whom he assigned the task of formulating methods and making recommendations for the instruction of feeble-minded children in the public schools of Paris. One member of this commission was the eminent psychologist, Alfred Binet, at that time Director of the Laboratory of Physiological Psychology at the Sorbonne. Binet was born in Nice, France, in 1857, and was educated to be a physician. His interests, however, early directed him into abnormal and child psychology, and it is in the latter field that he is best known. In 1895 Binet founded the journal *L'année Psychologique*, in which a number of his own studies and those of his students are published. As a direct outgrowth of his work as a member of the commission mentioned above, Binet in 1905, with the collaboration of Theophile Simon, published his first rough scale for measuring general intelligence by means of a set of tests. This first scale was followed by an enlarged and revised edition in 1908, and by a third and last revision in 1911, published shortly before Binet's untimely death.

It has often been said—and truly—that the development of the Binet-Simon Scale marks the real beginning of intelligence testing as we know it today. But it must not be thought on that account that Binet's scale is merely of historical interest. In spite of revision, criticism, discussion, and experimentation, it is today still the prototype of the best modern tests for measuring general intelligence, while Binet's conception of intelligence is in remarkably close agreement with the views of present-day psychologists. Less than ten years after the publication of its final revision, the Binet-Simon Scale was being extensively used in America, Canada, England, Australia, New Zealand, South Africa, Belgium, Germany, Switzerland, Italy, Russia, and China, and had been translated into Japanese and Turkish. Such widespread and immediate popularity dem-

onstrates clearly the great need for just such a device as Binet's intelligence tests provided.

2. DEVELOPMENT OF THE BINET-SIMON SCALE

How can we explain this remarkable response to Binet's simple set of tests? Perhaps the best way of answering this question is to consider just how Binet made his scale, how it differed from previous tests, and what it was designed to do. First, let us look at Binet's conception of intelligence, upon which the method and scale were built. Before Binet, mental tests had been devised chiefly with the intention of measuring rather narrow aspects or phases of mental ability, such as rote memory, attention-span, speed and accuracy of perception, discrimination in checking numbers and recognizing forms, speed and accuracy of voluntary movements, sensory acuity, and the like. Binet criticized tests of this sort as being too restricted in scope to be good measures of *general* ability. To measure intelligence, he said, we must tap the "higher mental processes," as revealed in reasoning, imagination, and judgment, since it is here that differences in intellectual ability are most likely to be manifested. Attention and adaptability, along with good judgment, were for Binet the most important contributors to intelligence. Intellect, he says, is compounded of "judgment, common sense, initiative, and the ability to adapt oneself." Again, he stresses "insight" into one's own capabilities (notoriously absent in the feeble-minded), the ability readily to adapt one's behavior to a definite end or goal, and persistence in sticking to a task once undertaken. Binet distinctly marked off intelligence from information acquired in school or in a good home, although he insisted that the intelligent person, unless removed from normal activities by untoward circumstances, will always know more than the unintelligent one.

Binet did not attempt to analyze intellect into its parts and then devise simple tests to measure these elements. On the contrary, he proposed, as a more promising approach, to estimate the level of intelligence by measuring the *combined* effects of attention, imagination, judgment, and reasoning, as shown in the performance of fairly complex tasks. Binet's first scale consisted of thirty carefully selected questions or problems. These were arranged in a rough order of difficulty which was determined by trying them out on about fifty normal children of different ages. The tests were not grouped according to age-level; Binet simply indicated how many tests a normal child of a given age should be expected to pass. In his second scale, that of 1908, the tests were for the first time arranged into age-groups, constituting the first *age-scale*. This scale contained from four to eight tests for each age-level from three to



ALFRED BINET

1857-1911



LEWIS M. Terman

1877-

thirteen years. Tests were assigned to specific age-groups in the following way. Whenever a test was passed by from 60 to 90 per cent of a given age-group, Binet considered it to be a fair test for children of that age. If all or nearly all of his five-year-olds failed in a given test, he regarded it as obviously too hard for children of that age; while if all of his ten-year-olds passed a test successfully it was clearly too easy for ten-year-olds. Binet's method of discovering at what age a given test should be located was somewhat rough, to be sure, but it possessed the virtue of being based upon experimental evidence rather than upon subjective impression.

The 1908 scale is of more than ordinary interest since here for the first time Binet employed the concept of "mental age." A child's mental age (MA) depends upon the number of tests in the scale which he can pass. If he does all of the tests assigned to the eight-year-old group successfully, he is said to have a mental age of 8 years, no matter what his chronological age (CA) is. If he happens to be five years old chronologically, he is, of course, advanced 3 years; if eight years old, just normal; and if ten years old, retarded 2 years. Description of performance in terms of mental age has proved to be quite useful in mental measurement. For one thing, "age scores" are more easily understood than are other kinds of scores by the non-psychologically trained person; and furthermore, such measures permit a quick and meaningful comparison of a child's years of mental growth with his years of physical growth, and of one child's mental level with that of another.

The 1911 scale represents three additional years of work with the 1908 scale and is the final form in which his tests were left by Binet. There are fifty-four tests in the 1911 scale, arranged in age-groups as follows:

Binet's 1911 Scale

Age III.

1. Points to nose, eyes, and mouth.
2. Repeats two digits.
3. Enumerates objects in a picture.
4. Gives family name.
5. Repeats a sentence of six syllables.

Age IV.

1. Gives own sex.
2. Names key, knife, and penny.
3. Repeats three digits.
4. Compares two lines.

Age V.

1. Compares two weights.
2. Copies a square.
3. Repeats a sentence of ten syllables.

4. Counts four pennies.
5. Unites the halves of a divided rectangle.

Age VI.

1. Distinguishes between morning and afternoon.
2. Defines familiar words in terms of use.
3. Copies a diamond.
4. Counts thirteen pennies.
5. Distinguishes pictures of ugly and pretty faces.

Age VII.

1. Shows right hand and left ear.
2. Describes a picture.
3. Executes three commands given simultaneously.
4. Counts the value of six sous, three of which are double.
5. Names four cardinal colors.

Age VIII.

1. Compares two objects from memory.
2. Counts from twenty to zero.
3. Notes omissions from pictures.
4. Gives day and date.
5. Repeats five digits.

Age IX.

1. Gives change from twenty sous.
2. Defines familiar words in terms superior to use.
3. Recognizes all of the (nine) pieces of money.
4. Names the months of the year in order.
5. Answers or comprehends "easy questions."

Age X.

1. Arranges five blocks in order of weight.
2. Copies two drawings from memory.
3. Criticizes absurd statements.
4. Answers or comprehends "difficult questions."
5. Uses three given words in not more than two sentences.

Age XII.

1. Resists suggestion as to length of lines.
2. Composes one sentence containing three given words.
3. Names sixty words in three minutes.
4. Defines three abstract words.
5. Discovers the sense of a disarranged sentence.

Age XV.

1. Repeats seven digits.
2. Finds three rimes for a given word in one minute.
3. Repeats a sentence of twenty-six syllables.
4. Interprets pictures.
5. Interprets given facts.

Adult.

1. Solves the paper-cutting test.
2. Rearranges a triangle in imagination.
3. Gives differences between pairs of abstract terms.
4. Gives three differences between a president and a king.
5. Gives the main thought of a selection which he has heard read.

Of the tests which had appeared in the 1908 edition, several were omitted from Binet's final revision of 1911 as a result of various criticisms. Binet left out those tests which had been found to depend chiefly upon specific information acquired in school, as well as such routine tests of common knowledge as knowing one's age and the days of the week. Several tests which were found to have been misplaced (being either too simple or too difficult) were eliminated or shifted up or down in the scale, and several new tests were introduced. In making this last revision, as in preparing his two earlier scales, Binet attempted to include only those tests which were not too difficult to administer, were relatively short, covered a wide range of mental activities, and did not depend directly upon specific school information. Such, in his opinion, were the attributes of a good intelligence test.

3. REVISIONS OF THE BINET SCALE IN AMERICA

Binet's tests were quickly taken up in America, where they were adapted and revised to fit American children and the conditions of American life. Goddard (1908) was the first to introduce Binet's tests into America. He translated the scale into English, made changes in the position and wording of certain tests, and used the scale in his work with the feeble-minded at the Vineland Training School in New Jersey. Several later revisions of the Binet Scale have appeared; one by Terman in 1916, a second by Terman and Merrill in 1937, and three by Kuhlmann in 1912, 1922, and 1939. In addition Yerkes, Bridges, and Hardwick (1915) and Herring (1922) published revisions in which the classification of the tests into age-groups was abandoned in favor of a "point scale" method. In point scales test items are arranged in order of difficulty, and the score in credits or points depends upon the number of tests passed. A point score may be translated into a coefficient of mental ability by dividing the child's score by the norm for his age (Yerkes); or it may be put into the more familiar MA or IQ (Herring).

The 1916 revision of the Binet-Simon Scale by L. M. Terman of Stanford University is called the "Stanford Revision," and often simply "Stanford-Binet." This examination represents a careful and thorough working-over of the old Binet Scale, and is by far the best known revision of Binet's tests. Terman found that tests in the original Binet scale were too few and too difficult at the upper age-levels; that many were misplaced in the scale; and that instructions for giving the tests were often indefinite. To insure uniformity of procedure in the final scale, one-half year was spent in training examiners to give the tests, and another half year in supervision of the testing. In all, thirty-one new tests

were tried out, the chief experimental group consisting of about 1,000 school children of average social status. As finally assembled, the Stanford Revision consisted of ninety tests, thirty-six more than Binet's 1911 Scale. Six tests and several alternates were placed in each age-group from three to ten; eight tests were placed at age twelve, six at age fourteen, six at "average adult" level (taken at sixteen), and six at the "superior adult" level (taken at eighteen). Terman used the expression *intelligence quotient*, or IQ, to express mental development or brightness.¹ The IQ is simply the ratio of the mental age to the chronological age. A child of eight with an MA of 8 years has an IQ of $8/8$ or 1.00; if the child's MA is 6 years, his IQ is $6/8$ or .75; if his MA is 10, his IQ is $10/8$, or 1.25. Mental age gives the intellectual *status* of the child, while the IQ tells how bright or how slow he is relative to the average child, whose IQ is always 1.00, or 100, as it is more commonly written. In the Stanford Revision, each of the six tests within a given age-group counts for 2 months of mental age (except at ages twelve, fourteen, sixteen, and eighteen, where each test has a greater value in months), so that mental ages of so many years plus so many months may be calculated from the scale.

The child of six years and six months, say, who is strictly average should pass *all* of the tests for the 6-year level and *below*, and three tests at the 7-year level. However, a strictly average child rarely exists except in textbooks, and so in almost every case there is some *scatter* in a child's record of passes: that is to say, he fails some of the tests below his true mental age and passes some of those above it. In the end these failures and successes tend to balance each other, so that the normal child of seven years and four months finally comes out with a mental age of 7 years and 4 months, and an IQ of 100. A superior child will, of course, go higher than his actual age, while a retarded child will find tests which he cannot do below his chronological age-level.

In 1937 a new revision of the "old" (1916) Stanford-Binet appeared. The new scale embodies the experience of the authors and of many other psychologists and is much more comprehensive than the old Stanford-Binet. Among the chief faults of the first revision the following should be mentioned: (1) the tests, and the subjects who served as the "standardization group," were too few at the extremes of the scale; (2) instructions for giving and scoring several of the tests were not precise enough, and some of the test items were particularly susceptible to coaching; (3) lack of a parallel or equivalent scale made it difficult to retest a child at a later period; (4) various technical faults and inaccuracies had been uncovered over the twenty years in which the old test had been used.

¹ The German psychologist, William Stern, seems to have been the first to employ the term *intelligence quotient* as a means of expressing mental development.

The new scale begins at age two; and from two to five there are tests at each one-half year level, viz., two and one-half, three, three and one-half, four, four and one-half, five. At these very low age-levels, an appeal is made to the interests of the very young child by the introduction of tests utilizing brightly colored pictures, beads, blocks, small toys, and the like.

Above age fourteen there is one set of tests for Average Adult level, and three complete sets of tests for the Superior Adult level. The Average Adult level represents an MA of 15 years, and is really the "ceiling" of the scale. Beyond this level, intelligence in the general population shows so little additional upward trend with age that MA's obtained from the Superior Adult level must be treated simply as numerical scores and not as true MA's. The location and scoring of the test items are based upon the responses of more than 3,000 children living in eleven states of the Union. These subjects are more representative of the American child population than were those used in the standardization of the old scale. The two equivalent forms of the new revision are called *L* and *M*. Each of these forms contains 129 separate items as against 90 items in the single form of the old revision. The new scale is being widely used. And its careful construction and many advantages in comprehensiveness and range have made it as popular as the old Stanford-Binet.

4. THE GROWTH CURVE OF GENERAL INTELLIGENCE

As we stated in the last section, the Average Adult level in general intelligence is placed by the Terman and Merrill Revision of Stanford-Binet at fifteen years. The view that intelligence matures or reaches the adult level at this relatively early age is often viewed with surprise not unmixed with suspicion. Many people will contend that intelligence surely continues to increase well into middle life, and they resent—when they do not deride—the idea that they are no smarter at forty than at fifteen. The confusion here is due in large part to a misunderstanding of what the psychologist means by *intelligence* and by *maturing*. By intelligence, it must be remembered, is meant ready adaptability to new situations, mental alertness, keenness, and ingenuity, *not* "knowledge" and "experience," which are products of these, and which usually do increase with age—at least beyond fifteen years. Generally speaking, the father of forty-five has more general—and more special—knowledge than his fifteen-year-old son; he can do many more things and has wider experience, though he may be no more alert nor readily adaptable (potentially intelligent) in the psychologist's sense. Careful and repeated measurements of the same and of different individuals over a period of years have shown that mental ability (insofar as it is measured by tests)

increases rapidly during the early years; then advances more and more slowly as the 'teens are reached, until somewhere between fourteen and sixteen years the *average* individual does as well as he will do at twenty or thirty years, or ever. At this comparatively early age, most people possess as much natural keenness and sheer native ability, apart from experience, as they will ever have.²

Curves showing the growth of intelligence in bright, average, and dull children are represented in Figure 39. The middle curve shows the course of average, the upper curve of superior, and the lower curve of

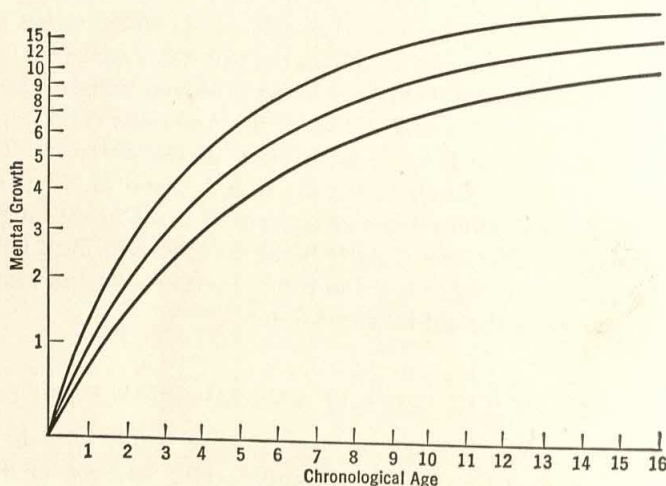


FIG. 39

MENTAL GROWTH CURVES FOR SUPERIOR, NORMAL, AND DULL CHILDREN FROM BIRTH TO MATURITY

inferior intellect. Note that average intelligence rises rapidly during the first four or five years; that from five to ten or eleven, growth, while evident, is considerably slower; and that from this point the curve gradually ceases to rise, becoming level somewhere between fourteen and sixteen years. The curve for superior intelligence rises more rapidly and goes higher than the average curve, probably not reaching its level until twenty years or even later. As shown by the lower curve, mental growth in the retarded (those of inferior intelligence) is distinctly slower than in the normal and superior. Also, in the retarded group mental growth reaches its maximum from two to three years earlier than in the normal, or average, group. Loss in general ability from twenty to about forty-five years appears to be relatively slight. In old age general intelligence as measured by tests wanes slowly.³

² See further Chapter 12, pp. 253-254.

³ See p. 254.

5. DISTRIBUTION OF INTELLIGENCE

The widespread application of intelligence tests to large numbers of children and adults has established quite clearly the range of intellect to be expected in the population at large. In statistical language, intelligence is said to be "distributed normally," that is, in accordance with the normal probability curve. Figure 40 shows a normal probability distribution or normal curve. Marked off on it are the percentages of the different grades or degrees of intelligence to be expected in a very large sample of people.

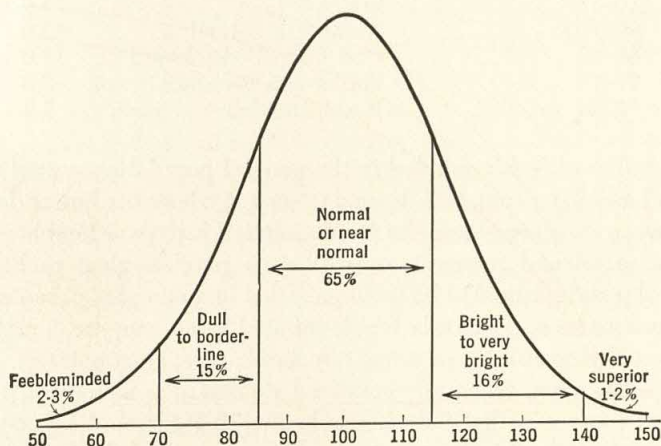


FIG. 40

DISTRIBUTION OF INTELLIGENCE (IQ's) IN THE GENERAL POPULATION ACCORDING TO THE 1937 REVISION OF THE STANFORD-BINET SCALE

A normal distribution of people with respect to any trait, such as height, speed of movement, or general ability, simply means that the majority possess a medium amount of the trait in question; that lesser and lesser amounts are possessed by fewer and fewer individuals; and that greater and greater amounts are possessed by steadily decreasing numbers. The normal distribution of a human trait or characteristic is well illustrated in the case of height. By far the greater number of men are close to five feet nine inches in height (the average); fewer men will be found two inches taller or two inches shorter than this average; still fewer four inches taller or four inches shorter; while the number of very tall (over seven feet) and very short (under five feet) men is exceedingly small indeed. What is true of height holds also for general intelligence, and probably for most mental traits. In terms of the revised Stanford-Binet about 65 per cent of the general population are classed as having average or

near average intelligence; about 15 to 16 per cent somewhat higher and the same per cent somewhat lower than average intelligence; while roughly 1 to 2 per cent are very superior and 2 to 3 per cent are mentally defective. (See Figure 40.)

A more detailed description of intellectual level in terms of IQ may be expressed as follows:

TABLE XXIV

IQ	Classification	%
140 and above	Very superior	1.5
120-139	Superior or very bright	11.0
110-119	Bright	18.0
90-109	Average or normal	48.0
80-89	Dull normal—backward	14.0
70-79	Border-line, very dull	5.0
0-69	Feeble-minded	2.5

The number of feeble-minded in the general population, usually put at between 1 and 3 per cent, will depend upon just where the line of demarcation between the normal and the feeble-minded is drawn. Feeble-mindedness is a social and economic as well as a psychological problem. An individual who is judged to be feeble-minded in a complex urban environment, because he is effectively feeble-minded there, may be "normal" (in the sense of being able to get along) in simple rural surroundings. In spite of this qualification, when a generalized statement is called for, the best opinion at present is that individuals below 70 IQ should be considered mentally defective, at least potentially. Three classes of the feeble-minded are recognized, idiots, imbeciles, and morons. Idiots, who are the lowest in intellect, have IQ's (when measurable) under 25; and mental ages under 3 years. Idiots use only a few simple words. They cannot learn to wash or dress themselves and must be cared for in an institution or in the home. Imbeciles range in IQ from 25 to about 50, with mental ages of from 3 to 7 years. Imbeciles learn to talk and can even read very simple prose. But they lack foresight and planning ability and can perform only the simplest routine tasks and these under supervision. Morons, the highest level of feeble-mindedness, have IQ's from 50 to 70; their mental ages range from about 8 to 10 years. Morons can often be taught many useful tasks: sewing, washing dishes, housework, and simple mechanical jobs. When not closely supervised, however, the moron may become a serious social problem. Under the influence of vicious persons the moron boy is likely to steal or commit even more serious crimes; while the moron girl is often led into sex delinquency and prostitution.

Worth mentioning in the present connection is a contribution made to

theoretical psychology by the Binet and other mental tests. This is the experimental finding that the feeble-minded are not a distinct species separated sharply from the normal as, for instance, animals are separated from man. There is no sharp cleavage between normal and feeble-minded insofar as performance on intelligence tests is concerned, but rather a gradual progression from one to the other. The difference between the two groups is *quantitative* rather than *qualitative*; a matter of more available and more complex behavior resources rather than of a different *kind* of intellectual activity.

Binet's original tests were constructed with the expressed purpose of identifying the feeble-minded, and for a long time this was regarded as the chief function of the general intelligence test. Probably a good deal of the suspicion with which intelligence tests were (and, to a degree, still are) regarded is a reflection of this view. The expression "to submit" to an intelligence test describes this attitude exactly. With the growing importance of mental tests as a means of locating superior children who should be rapidly advanced or given special attention, distrust of intelligence measurement is, fortunately, on the wane. Terman's studies of 1,000 children with IQ's of 130 and above have shown that these superior children tend to be above the average in height, weight, and general health, as well as in looks, and in social and emotional maturity (see further p. 235). This flatly contradicts the old idea—still prevalent—that precocious children are poorly adjusted socially, tend to be puny and undersized, and are apt to die at an early age! The parents of the very bright children tested by Terman were for the most part from the professional and semiprofessional classes (80 per cent), only a small fraction (6 per cent) coming from the semiskilled and laboring classes. One fourth of the children had at least one parent who was a college graduate, while the average schooling of both parents was about twelve grades—twice that of the general adult population. These results indicate that both the heredity and the environment of these bright children were distinctly above average.

6. INTELLIGENCE TESTING OF ADULTS

The Stanford-Binet is intended for use with children, and test items were selected primarily with the interests of youngsters in mind. Many psychologists had long felt the need for an individual intelligence test with more "adult" items for use with older subjects. Adults do not like questions which they regard as childish and somewhat silly—even when they cannot give the correct answers. Asking a moron of forty, for example, to copy a bead chain or what he should do when it appears that

he will be late for school does not increase his coöperation with the examiner nor his respect for the test. The need for an adult scale was remedied by the publication in 1939 of the Wechsler-Bellevue Intelligence Scale. This test is intended for adolescents and adults but it may be used with children as young as ten years of age.⁴ The Wechsler-Bellevue Scale includes ten sub-tests plus one alternate. These eleven sub-tests are designed to measure the following:

- 1) Common information and knowledge
 - 2) Understanding and comprehension as shown in sensible answers to questions
 - 3) Memory span for series of numbers which are to be repeated forward or backward
 - 4) Knowledge of similarities as between word pairs
 - 5) Reasoning shown in the mental solution of arithmetic problems presented orally by the examiner
 - 6) Location of missing parts in a series of pictures
 - 7) Arrangement of sets of pictures in order so that they tell a story
 - 8) Correct assembly of parts in three formboards to make a hand, a man, and a face
 - 9) Copying of designs by fitting together differently colored cubes
 - 10) Associative learning (digit-symbol test) in which numbers are marked with symbols in accordance with a key
- Alternate:* A vocabulary test consisting of words (graded in difficulty) to be defined orally

The first five tests plus vocabulary constitute a *verbal* scale, the second five a *performance* scale. The full scale includes both the verbal and performance tests. Like the Stanford-Binet, the Wechsler-Bellevue yields an overall index of intelligence. Knowing an individual's chronological age and the sum of his scores on the separate sub-tests we can determine his IQ by means of appropriate tables. This scale does not yield a MA as the concept of mental age is really meaningless when used in reference to adults (see p. 221). The Wechsler-Bellevue IQ, therefore, is not a ratio (MA/CA) as is the Stanford-Binet IQ. Instead it is a score which is given a consistent meaning in the following way. The median test score for each age level is set at 100 IQ. A man of forty-five who achieves a score of 81 has an IQ of 100 and is just average for his age group; while a young man of twenty-five must achieve a score of 95 to earn an IQ of 100—be average for his group. Scores above and below the medians are assigned IQ's in accordance with their percentage distances from the middle point. A man who earns a total score of 80 on the full scale, for

⁴ An adaptation of the Wechsler-Bellevue Scale—the Wechsler Intelligence Scale for Children—was published in 1949. This examination employs essentially the same sub-test forms found in the Wechsler-Bellevue, adapted for use with children from five to fifteen years old.

example, has a Wechsler-Bellevue IQ of 87 if he is twenty years old, 92 if he is thirty, 97 if forty, and 101 if fifty. These figures show clearly that performance on the average declines steadily with advancing age. Figure 41 gives an idea of the extent of this performance loss. From about twenty years onward the fall-off is steady and progressive up to 65 (see p. 254).

After the IQ has been computed from the ten tests of the Wechsler-Bellevue Scale, the clinical psychologist usually examines the range of scores made on the various sub-tests. A profile like that shown in Figure 42 shows graphically the "scatter" or discrepancies on the Wechsler-Bellevue test. Divergencies from the average or norm may often be large. The subject in Figure 42, for instance, achieves higher scores on the per-

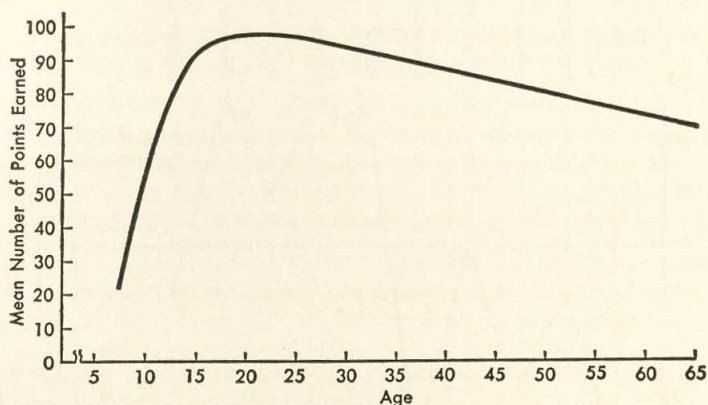


FIG. 41

CHANGES IN MENTAL ABILITY WITH AGE

(From Wechsler, *Measurement of Adult Intelligence*, 3rd ed., 1944.
The Williams & Wilkins Company.)

formance than on the verbal tests. A second subject may rate uniformly low in the "planning" tests (comprehension, picture arrangement, object assembly) and higher on the tests of routine knowledge (information, vocabulary); a third lower in digit span (a test which requires close attention) than in comprehension or similarities where the attention factor is not so crucial; and still a fourth may exhibit wide swings from test to test without revealing *any* clear pattern. Indices of scatter must be employed with caution, as the various sub-tests are not long enough to insure highly stable part scores. When supplemented by other information, however, discrepancies among sub-test scores will aid in the better understanding of a subject's vocational and educational history and provide as well a basis for further guidance.

The Army Individual Test of General Mental Ability (1944) is another intelligence test designed specifically for use with adults (soldiers in

World War II). This mental examination consists of six subtests, three verbal and three performance. The three verbal tests are: (1) similarities-differences (the examinee is to tell in what way pairs of words are alike or different); (2) story-memory (a paragraph is read to the examinee who must repeat it and answer questions); (3) digit span (memory for numbers presented orally).

The three performance tests are (1) shoulder patches (examinee must reproduce a colored design from colored patches); (2) trail making (lines

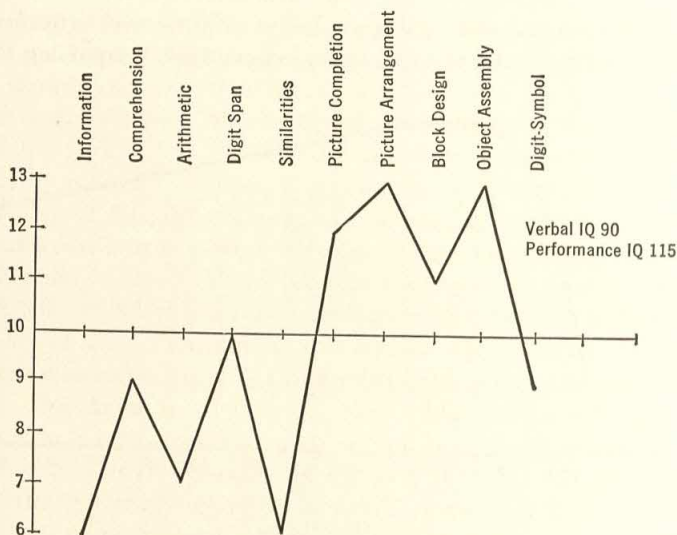


FIG. 42
PROFILE OF A SUBJECT ON THE FULL SCALE OF THE WECHSLER-BELLEVUE
INTELLIGENCE SCALE

must be drawn from number to number or letter to letter in sequence); (3) cube assembly (the examinee must duplicate various designs by making arrangements of 24 one-inch cubes).

The sum of the six subtest scores gives the final test score. The MA and the IQ are not used. This individual test is interesting, well constructed, and economical of time (forty minutes are required for the examination). It provides a good alternate test to the Wechsler-Bellevue Scale.

7. CRITICISMS OF THE INDIVIDUAL INTELLIGENCE TEST

Among the criticisms which from time to time have been leveled against the individual intelligence test, probably the most common is that a subject's performance depends more upon training and social surroundings

than upon native ability.⁵ For this reason, it is said, the child from an educated and cultured home has a distinct advantage over one not so favored. This is quite generally true, provided the cultural gap is wide, and has never been disputed by psychologists. Native capacity is of little value per se unless the environment is conducive to normal development. An American white child brought up from birth among Eskimos would undoubtedly test as feeble-minded on the Binet Scale, no matter how great his potential ability. Or, to take a scarcely less obvious illustration, it is clearly unfair to compare children from the slums of New York City with children from cultured and well-to-do homes. Experimental studies have shown that environment has a differential effect upon mental age and IQ. Adoption into a good (cultured) home, for example, may raise a child's IQ 10 points or even more; while adoption into a very poor home may depress the IQ by as much as 10 points. Cyril Burt, an English psychologist, found marked differences in mental test performance in favor of children of superior social status in London. The same investigator found also that children from superior schools were on the average 1 to 2 years in mental age ahead of those from inferior schools. In the United States children in isolated rural communities often rank considerably below city children in mental age. In one study, younger rural children (seven to nine years old) were found to be retarded from 1 to 1½ years, while older children (thirteen to fifteen years old) were retarded by as much as 5 or even 6 years. Probably the reason for the greater backwardness of the older children lies in the fact that poor social-economic status—even when it remains substantially constant—becomes progressively more restrictive as children grow older. Results like these emphasize the importance of considering the child's environment when interpreting his mental test rating or judging his mental development.

Psychologists have been well aware that intelligence tests do not measure "innate" ability completely divorced from the influences of environment. They believe that a comparison of the mental ages of two children is valid *only* when both children have had approximately the same schooling and the same common background of experience. But when these conditions *are* satisfied, test ratings are fair measures of comparative ability on the reasonable assumption that bright and dull children, when exposed to the same facts, will learn according to their native endowment. Binet's tests were constructed with the avowed purpose of reducing to a minimum the influences of special training. Unless differences in social and economic status are great, therefore, or the deprivations arising from lack of normal contacts with the environment serious, mental ages are close estimates of "real" ability. To be sure, answers in a mental test often

⁵ In this connection, read the discussion in Chapter 12, pp. 261-262.

require the use of words and these are obviously learned. But it is through language responses that the subject reveals the ability to see relations, interpret meanings (in a picture, for instance), give sensible definitions, detect absurdities or incongruities, and solve problems demanding thought, reasoning, and judgment. Such performances are the real indices of intellect. In short, although the individual intelligence test draws heavily upon language, it demands a high degree of sagacity, cleverness, and mental alertness, rather than the ability simply to reproduce parrot-like facts learned by rote.

Another criticism sometimes brought against the individual intelligence test is that it fails to detect personality and character defects, such as bad temper, laziness, and abnormal sex habits, which are as important as intellect in daily living. Also, it is sometimes charged that mental tests fail to discover exceptional gifts in music, art, or mechanics. The reply to this is that the individual mental test is designed to gauge the intellectual level at which the subject habitually functions, not to reveal character traits or to discover exceptional aptitudes. At the same time, from careful observation of an examinee's behavior the trained psychologist may gain many clues as to his emotional and social maturity. The timid, insecure child who asks again and again if he is right; the stubborn, negative child who is defiant or aggressive; the spoiled child who gives up and pouts when he cannot get the answer—all these reveal their personalities in reacting to test situations. The mental test record may yield valuable clues, therefore, even when one is interested primarily in problems of emotional maladjustment. Time and again juvenile courts and children's clinics have demonstrated an intimate connection between crime and immorality on the one hand, and low-grade intelligence on the other. The plan to be followed in dealing with character defects and delinquencies must depend to a considerable degree upon the intellectual level of the offender. Knowing the mental age rating simplifies the search for causes, even though the problem turns out to be chiefly an emotional one.

Verbal intelligence tests cannot be used effectively with those who do not speak English, nor can they be readily employed with the deaf, or with those who stutter, stammer, or have other language handicaps. In many cases, therefore, language tests must be supplemented (or replaced) by performance tests. These tests require the examinee to fit blocks into a board (form-board tests), reconstruct puzzle pictures, learn mazes, and, in general, carry on mental activities which demand ingenuity and insight, with a minimum of language.⁶

In closing this section a word of warning is in order regarding the administration of the Stanford-Binet, or of any individual mental test.

⁶ See also pages 137-139.

Because the tests seem so easy to give, many people mistakenly believe that they can administer them without previous supervision in test-giving or training in psychology. The unfortunate part of it is that they *can* give the tests after a fashion, but the results will almost certainly be worthless and will often be downright misleading. Such "testers" forget that the object of a mental examination is to see whether a subject can do certain things under *carefully prescribed conditions*, not whether he can do them when given plenty of time, and often plenty of suggestion and prodding as well.

The individual intelligence scale is a valid measure of general intelligence when—and only when—the tests are given and scored in a standard manner; when the examinee's confidence and coöperation have been gained and held; when the subject comes from a "normal" environment, and labors under no marked physical handicaps. The individual intelligence examination should always be supplemented by careful observation in order to detect peculiarities or abnormalities which might affect the examinee's score. The personal history, medical record, and habits of industry of the subject should be recorded, as well as his social and moral behavior. When the subject is a child, play habits and the personal traits exhibited in school should be carefully noted. A year's graduate study in psychology plus at least six months' practical experience in giving mental tests under expert supervision is a minimum requirement for a trained examiner.

8. THE INHERITANCE OF INTELLIGENCE

The question of the constancy or stability of an individual's general level of intelligence throughout life is bound up with the prior consideration of whether intelligence is inborn and native, and hence largely if not entirely determined at birth; or whether intelligence is only loosely native in the sense that all behavior is potentially native, and hence is highly susceptible to training. There is considerable evidence in favor of the view that intelligence is inborn—inherited in much the same manner as are physical characteristics.

In the first place, we have the facts of everyday observation. Bright boys and bright girls tend to grow into bright men and women, while the history of stupid adults reveals more often than not a record of stupidity in childhood. Nearly everyone can cite instances from his own experience of dull children, who, in spite of every advantage socially and educationally, grew into dull and mediocre men and women; while we all know of bright children, who, despite marked disadvantages, insisted upon remaining bright and becoming successful adults. Among children who

grow up in the same community, attend the same schools, play together, see the same movies, and, all in all, possess about the same opportunities, some will learn more rapidly and progress faster than others. Such an outcome must be largely the resultant of native differences in endowment.

There are exceptions, of course. Disease, deafness, poor eyesight, unkind treatment, a restrictive or vicious environment—all of these serve to complicate the nature-nurture problem. But when such factors are allowed for, the fact still remains that some people are natively better equipped mentally than others.

The biographies of men of genius show that exceptional gifts usually appear early in childhood. A famous illustration is Sir Francis Galton, who could read when he was two and one-half years old. The day before his fifth birthday he wrote the following note to his sister.

My dear Adele, I am 4 years old and can read any English book. I can say all of the Latin substantives and adjectives and active verbs besides 52 lines of Latin poetry. I can cast up any sum in addition and multiply by 2, 3, 4, 5, 6, 8, 10. I can also say the pence table. I can read French a little and I know the clock. Francis Galton. Febuary (sic) 15, 1827.

Terman, after studying Galton's biography and later career, estimates that his IQ must have been close to 200. Voltaire began to read at three, and at twelve wrote a tragedy. Sir Isaac Newton, as a child, was "constantly occupied during his play hours" in devising all sorts of contrivances and machines, especially water-clocks and kites. At seventeen, Goethe had done extensive reading in German, French, Latin, and Hebrew; he knew the history of the chief European countries in detail; he played the piano and the flute, and was considered a promising art student. According to Cox, who has made an exhaustive study of the childhood of eminent men, Goethe's IQ must have been between 185 and 200. There are many other illustrations of the precocity shown in childhood by great men. Macaulay could read at four, and at eight wrote a "treatise to convert the natives of Malabar to Christianity"; Jonathan Edwards at twelve wrote a paper on spiders which actually increased the scientific knowledge of the subject; Sir Walter Scott (estimated IQ 150) and John Stuart Mill (estimated IQ 190) were child prodigies. The childhood histories of Francis Bacon, Descartes, Spinoza, and many others give early promise of later greatness.

Not all geniuses, however, are recognized as precocious in youth. Charles Darwin (estimated IQ 135), for instance, was considered rather dull by his teacher, partly, no doubt, because he carried insects and small animals around in his pockets, oftentimes disturbing the serenity of the classroom.

Napoleon stood forty-second in his class in military school. Of David Hume it was said by his mother that he was good-natured but "uncommon wakeminded." Thomas A. Edison was usually at the foot of his class in school, and was considered "addled" by his teacher despite the fact that he had read Gibbon's *Decline and Fall of the Roman Empire*, Hume's *History of England*, and Burton's *Anatomy of Melancholy* before he was twelve. It is hardly probable that these illustrations are exceptions to the rule, "bright child, bright adult." Many brilliant children are hopelessly misunderstood by their parents as well as by their teachers. Uninterested in the simple facts taught in school, they often neglect their lessons concerning themselves with matters far beyond the capacity or interests of the ordinary child. Bright children, too, are frequently mischievous ("problem cases") because, quickly grasping what is taught, they idle or play instead of dutifully paying attention and carrying out assigned tasks. Intelligence tests are useful in discovering and providing for the exceptional child. Special classes, plus extended and enriched programs, offer opportunities for initiative and creative endeavor, so that the energy of the bright child is usefully conserved and applied.

At the other end of the intelligence scale, there is much evidence—from the way in which feeble-mindedness tends to run in families—that intelligence is native and inherited. Authorities estimate that perhaps 80 per cent of all feeble-mindedness is native and inherited, the remaining 20 per cent being due to disease or injury at birth or in early childhood (Penrose, Tredgold). Normal parents who come from families in which there is no mental defect rarely have a feeble-minded child. If one parent is normal and the other feeble-minded, some of the children are likely to be feeble-minded or dull. If both of the parents are feeble-minded, *all* of their children will be feeble-minded or at least low-grade.

Many studies have shown that feeble-minded tend to mate with feeble-minded, thus passing on the defective strain from one generation to the next. Unfortunately, however, family histories, though interesting and suggestive, do not enable us to separate out clearly the distinctive contributions of heredity from those of environment. Dull and ineffective people tend to gravitate to mean surroundings, and superior people to good surroundings—on the general principle that water seeks its level.

In spite of its limitations, it will be enlightening to review one of the best-known family histories, namely, Goddard's account of five generations of the Kallikak family. During the American Revolution, Martin Kallikak (the name is fictitious), a man believed to be of normal mentality, had an illegitimate son by a girl known to be defective mentally. From this son have come 480 descendants. Of these, 143 were judged to be feeble-minded, 46 probably normal, and the rest doubtful. Among the

total lot, 36 were illegitimate, 33 were sexually immoral (mostly prostitutes), 24 were drunkards, 3 were epileptics, 82 died in infancy, 3 were criminals, and 8 were keepers of houses of ill-fame. Of this family tree Thorndike remarks that it constitutes "a horrid array of human incompetence." After the Revolution, Martin Kallikak married a normal woman of Quaker stock. Goddard was able to trace 496 descendants of this union, and the records of these offspring furnish an interesting parallel to the other line. Of these legitimate progeny all, except possibly two, were normal mentally and morally, and several were evidently of superior intellect. In this group we find lawyers, physicians, governors, professors, and college presidents. Two groups differing more sharply in achievement and social value can scarcely be conceived.

The data of the Kallikak family are striking and dramatic; but they do not offer conclusive evidence of the inheritance of feeble-mindedness. In the first place, the environments of the two parallel lines did not provide anything approaching equal opportunity. Again, the judgments of feeble-mindedness made by Goddard's field workers were often based upon dubious evidence, amounting to little more than hearsay. Further, the heredity of the two lines was alike as regards the common paternal ancestor, Martin, so that the case for heredity rests entirely upon the maternal lines. While not so crucial an argument for the inheritance of feeble-mindedness as Goddard thought, this investigation is nonetheless convincing. The Kallikak study shows what may be expected when initially good and bad strains are placed in good and bad surroundings. Environment alone could hardly have produced the enormous differences found if the two family groups had not possessed markedly different capacities initially.

Better evidence of the influence of heredity in determining level of intelligence has come from the study of mental resemblances among members of the same family, and from comparisons of adopted and of own children. Twins are more alike than siblings (ordinary brothers and sisters) in traits little affected by training; and also in traits upon which the school concentrates its influence. In one of the best-controlled studies of adopted children, it was found that a considerably higher relationship existed between the intelligence of children and their own parents than between adopted children and their foster parents. This result is significant because the age of adoption was low (six months or less), the group large (194), and when examined the children had been in foster homes long enough (four to ten years) for the effects of home environment to be felt. The degree of resemblance between members of a family in mental traits is as high or higher than is their resemblance in physical traits such as height, weight, hair-color, and eye-color. Since the physical

resemblances of twins and of brothers and sisters can be attributed to native factors, it follows that mental resemblances must, in part at least, also have a genetic basis.

9. THE CONSTANCY OF THE IQ

If intelligence is determined largely by native factors, the IQ—assuming it to be a fair measure of mental ability—should remain constant throughout life. Many studies have shown that the IQ is substantially constant over the age-range covered by the Stanford-Binet Scale. It must be remembered, however, that even the best mental test is not a perfect measuring instrument. Hence, we cannot expect repeated testings of the same individual on the Stanford-Binet or the Wechsler-Bellevue to give *exactly* the same result. The amount of change which must be allowed for from test to retest is expressed by the “probable error” of a score. In the 1937 Stanford-Binet, the probable error of a calculated IQ is 3 to 4 points; in the Wechsler-Bellevue it is 5 points. A change in an IQ from 95 to 98 or 92 (or even a greater change) does not imply, then, any necessary change in a person’s mental ability. He is still within the normal range. Repetitions of the Stanford-Binet upon the same child after an interval of one year or less will seldom show an IQ variation of more than 3 or 4 points. When the time interval between test and retest is longer than a year, variation in IQ may be as much as 5 or even 10 points, but this much change is observed very infrequently.⁷

There are other reasons (besides errors of measurement) why a child’s IQ may fluctuate from time to time without leading us to suspect any real change in the child’s mental status. It may be of interest to list some of the more important of these factors.

(1) *Physical Causes*. These are of primary concern. On the first administration of a test, deafness, poor eyesight, and other sensory defects may pass unnoticed, thus depressing the score. Malnourished children brought to institutions from bad homes will frequently show improvement in measured intelligence after a few months of kind treatment, good food, and medical attention, the latter leading, say, to the removal of tonsils and adenoids. Children who come from poor social and economic levels, especially if handicapped by sensory defects, chronic illness and deformities, often show marked increases in IQ after attending a school in which the training is individual and personal.

⁷ Several studies have contended that the IQ can be increased by proper training to a dramatic degree (refs. 27, 40). The experimental and statistical methods employed in these studies have been sharply challenged by competent authorities, however, and their conclusions cannot be accepted without qualification.

(2) *Emotional Causes*. One can hardly expect to get a fair measure of intelligence when the child is badly frightened, angry, or otherwise emotionally disturbed. Children who are badly spoiled, overindulged, pampered and dependent, make poor subjects for a mental test.

(3) *Examiner*. The "personal equation" of the examiner is extremely important. If an examiner is poorly trained and inexperienced, possesses a harsh voice or peculiarities of physique or appearance, or if he is timid, supercilious, or disagreeable in manner, he may get very little coöperation or response. Many children have been observed to respond with enthusiasm to a task presented by one examiner, and with indifference or distaste to the same task presented by another.

(4) *Testing Conditions*. If the examination room is small, bare, or over-decorated, the temperature too high or too low, the lighting bad or the air stale, interruptions or distractions (opening or shutting doors) frequent, the examination is almost inevitably affected.

(5) *Tests*. Only those tests can be expected to yield consistent and valid results from one time to another which have been carefully constructed, have been standardized upon large groups, are expertly administered, and have stable norms.

10. IQ AND SCHOOL SUCCESS

The brightest youngster is not always the most successful in school. As every one knows, appearance, manner, docility, industry, personality and character definitely influence school grades and often compensate for lack of mental ability. But the level of intelligence sets the upper limit to a child's progress, and a knowledge of MA and IQ is of great value in educational guidance. The educational expectation of children of different mental levels may be set down roughly as follows:

(a) Children under 70 IQ, when in school at all, remain in the first grade until ten or eleven years old. A very few may reach the fifth grade by fourteen or fifteen; seldom if ever will a child as dull as this go beyond the third or fourth grade.

(b) Mentally deficient to dull children (70 to 85 IQ) will ordinarily be one to two years over age for their grade. The eighth grade is the upper limit for many of these children; most of them will not get beyond the fifth or sixth.

(c) Children with IQ's of 85 to 115 make up the bulk of the elementary school population. Most of these children will finish elementary school and many high school; but those below 100 IQ will have difficulty in graduating from a good high school.

(d) Children above 115 IQ are normal or one to two years accelerated

for their grade. These children are good college material. If they possess the necessary personality and character traits, they should be successful in business or professional life.

Several illustrations will show how a knowledge of intelligence level is helpful in educational guidance.⁸

C.P. Boy, age 10-2; MA, 7-11; IQ 78, border-line, dull. Son of skilled laborer, one of 11 children. Is in third grade but cannot do work. Is capable in games. Will probably never be able to do school work beyond fifth or sixth grade. Except for extreme stubbornness his social development is fairly normal.

G.G. Girl, age 12-4; MA 10-10; IQ 82, dull. From average home. Has had excellent educational advantages and has no physical handicap. Is doing poor work in fifth grade. Appearance, play life, attitude toward other children normal. Will probably never go beyond eighth grade.

M.P. Girl, age 8-10; MA 11-1; IQ 130, superior. Father unusually intelligent house painter. Home above average. M.P. now in fourth grade, marks "excellent." Good health, good social and moral traits. This child will be ready for high school by the age of 12.

E.M. Boy, age 6-11; MA 10; IQ 145, superior. Father university professor. E.M. now 7-8 is in fourth grade. Learned to read by himself at 5, and shortly afterward learned all the capitals of the states in the Union. Entered school in first grade and by the end of the day had been moved up to the fourth grade. Personality and health excellent. E.M. should be ready for college at the age of 15.

BIBLIOGRAPHY

1. Asher, E. J. The inadequacy of current intelligence tests for testing Kentucky mountain children. *J. genet. Psychol.*, 1935, 46, 480-486.
2. Binet, A., & Simon, Th. *L'intelligence des imbéciles*. *L'Année Psychologique*, 1909, 15, 1-147.
3. Binet, A., & Simon, Th. The development of intelligence in children. Translated by E. S. Kite, Baltimore: The Williams and Wilkins Company, 1916.
4. Brown, R. R. Time-interval and constancy of IQ. *J. educ. Psychol.*, 1933, 24, 81-96.
5. Burks, B. S. Twenty-seventh Yearbook of the National Society for the Study of Education, 1928, Part I, 219-316.
6. Burt, C. Mental and scholastic tests. London County Council, 1921, 190-198.
7. Cattell, P. Constant changes in the Stanford-Binet IQ. *J. educ. Psychol.*, 1931, 22, 544-550.
8. Cox, C. M. The early mental traits of three hundred geniuses. In Terman, L. M., *Genetic studies of genius*. Stanford University Press, 1926, Vol. II.
9. Davis, A. *Social-class influences upon learning*. Cambridge: Harvard University Press, 1948.
10. Downey, J. E. *The will-temperament and its testing*. Yonkers, New York: World Book Company, 1922.
11. Garrett, H. E., & Schneek, M. R. *Psychological tests, methods, and results*. New York: Harper & Brothers, 1933, Part I.

⁸ Adapted with some abridgement from L. M. Terman.

12. Goddard, H. H. *The Kallikak family*. New York: The Macmillan Company, 1912.
13. Hollingworth, L. S. *The psychology of subnormal children*. New York: The Macmillan Company, 1920.
14. Hollingworth, L. S. *Gifted children, their nature and nurture*. New York: The Macmillan Company, 1926.
15. *Intelligence: Its nature and nurture*. Thirty-ninth Yearbook of the National Society for the Study of Education, 1940, Parts I and II.
16. Jones, H. E., & Conrad, H. S. The growth and decline of intelligence. *Genet. Psychol. Monog.*, 1933, 13, 223-298.
17. Kelley, T. L., Ruch, G. M., & Terman, L. M. New Stanford achievement test; advanced examination, directions for scoring. 1931, 7-9.
18. Leahy, A. M. Nature-nurture and intelligence. *Genet. Psychol. Monog.*, 1935, 17, 236-308.
19. Louttit, C. M. *Clinical psychology*. New York: Harper & Brothers, 1936, especially chaps. 4, 5, 6.
20. McNemar, Q. A critical examination of the University of Iowa studies of environmental influences upon the IQ. *Psychol. Bull.*, 1940, 37, 63-92.
21. Merrill, M. A. The significance of IQ's on the revised Stanford-Binet scales. *J. educ. Psychol.*, 1938, 29, 641-651.
22. Miles, W. R. Psychological aspects of ageing. Chap. 20 in *Problems of ageing*, edited by E. V. Cowdry. Baltimore: The Williams & Wilkins Company, 1939.
23. Penrose, L. S. *Mental defect*. London: Sidgwick & Jackson, 1933.
24. Pintner, R. *Intelligence testing, methods, and results*. New York: Henry Holt and Company, 1931, Chaps. I, II, III.
25. Rogers, C. R. *The clinical treatment of the problem child*. Boston: Houghton Mifflin Company, 1939.
26. Sandiford, P. *Foundations of educational psychology*. New York: Longmans, Green & Company, 1938.
27. Schmidt, B. Changes in personal, social and intellectual behavior of children originally classified as feeble-minded. *Psychol. Monog.*, 1946, 60, 5.
28. Seashore, C. E. *Psychology of music*. New York: McGraw-Hill Book Company, Inc., 1938.
29. Staff, Personnel Research Section, Classification and Replacement Branch, A.G.O., The new army individual intelligence test of general mental ability, *Psychol. Bull.*, 1944, 41, 532-538.
30. Terman, L. M. *The measurement of intelligence*. Boston: Houghton Mifflin Company, 1916.
31. Terman, L. M. The intelligence quotient of Francis Galton in childhood. *Amer. J. Psychol.*, 1917, 28, 209-215.
32. Terman, L. M., et al. *Genetic studies of genius*. Stanford University Press, 1925, Vol. I.
33. Terman, L. M., & Merrill, M. A. *Measuring intelligence*. Boston: Houghton Mifflin Company, 1937.
34. Thorndike, E. L., et al. *The measurement of intelligence*. New York: Teachers College, Columbia University, 1927.
35. Thorndike, E. L., et al. *Adult learning*. New York: The Macmillan Company, 1928.

36. Tredgold, A. F. Mental deficiency. New York: William Wood and Co., 1929.
37. Wechsler, D. The measurement of adult intelligence, 3rd ed. Baltimore: The Williams & Wilkins Company, 1944.
38. Wechsler, D. Wechsler intelligence scale for children. New York: The Psychological Corporation, 1949.
39. Weisenberg, T., Roe, A., & McBride, K. E. Adult intelligence. New York: Commonwealth Fund, 1936, 15-32.
40. Wellman, B. L., Skeels, H. M., & Skodak, M. Review of McNemar's critical examination of the Iowa studies, Psychol. Bull, 1940, 37, 93-111.

12

THE ARMY TESTS: ALPHA, BETA, AND AGCT

I. THE ARMY TESTS OF GENERAL INTELLIGENCE

(1) *Army Alpha and Army Beta*

THE WIDESPREAD use of group intelligence tests during World War I for the purpose of classifying quickly large numbers of soldiers constituted psychology's first great experiment in human engineering. During the years 1917-1918, intelligence examinations were given to approximately 1,750,000 men. As a direct result about 8,000 men were recommended for discharge because of limited ability; something like 10,000 were assigned to labor battalions or to other services requiring only a low level of intelligence; and about 10,000 were sent to special development battalions for observation and further training. Nearly one third of all men examined were unable to read or write, or else did so too poorly to be classed as literate. To these was given another examination prepared for illiterates.

The Army mental tests were not, as is sometimes supposed, a series of riddles, puzzles, and other "stunts" thrown together with little purpose or design. On the contrary, these tests were carefully selected and methodically assembled in accordance with the best principles of test making then available. In April 1917 a committee was appointed by the American Psychological Association to which was entrusted the task of preparing an adequate examination for measuring the intellectual level of large groups of men at the same time. This committee decided that an intelligence test for use with soldiers should meet the following requirements: (1) It should as far as possible be independent of specific school information, since the aim of the test was to measure a man's native wit rather than the extent of his formal school training; (2) it should be steeply graded in difficulty, hard enough to tax the high-grade men and easy enough to measure those of lesser ability; (3) scoring should be simple, rapid, and objective, so that little would be left to the personal judgment of the scorer; (4) answers should be as brief as possible in order to eliminate speed of writing as an important factor in determining the score;

(5) several different forms approximately equal in difficulty should be drawn up to prevent coaching on a particular form; (6) the test should utilize material interesting and varied enough to keep the man "on the job." How well these standards were met we shall see in later sections.

Two tests were devised, one for men who could read and write and one for those who could do neither (or else do so very poorly). The test for literates was called Alpha, that for illiterates, Beta. In preliminary trials Alpha was tried out on elementary and high school children; students in schools, colleges, and officer training camps; more than 5,000 enlisted men; and the inmates of various institutions for the feeble-minded. The *validity* of the examination, that is, its claim to be a measure of intellect, was checked against various criteria: among the students and feeble-minded by such measures of aptitude as school grades, teachers' estimates of ability, and the Stanford-Binet; among soldiers by officers' ratings for ability, rank attained, aptitude shown in training, previous civilian accomplishments, and the like. The correlations between Alpha and various criteria ranged from .50 to .80, so that the test was in a statistical sense a good measure of general ability as defined. Results for Beta were comparable to those for Alpha. The *reliability* or consistency with which Alpha and Beta measured ability proved to be satisfactory. On taking a second form of Alpha, for example, a soldier's score rarely deviated more than four or five points from his first rating. An obtained Alpha score was, therefore, a relatively stable measure of performance.

(2) *The Army General Classification Test (AGCT)*

During World War II (i.e., from about 1940 to 1945), a new test, the Army General Classification Test (AGCT) was administered to some 12,000,000 soldiers. Like Army Alpha, AGCT is a verbal group intelligence test, intended for use with men who can read and write English. Illiterates and those who fell in the lowest 5 to 10 per cent of AGCT scores were given individual non-language tests—somewhat like Army Beta. AGCT was an improvement over Army Alpha in several respects. For one thing, instead of being grouped into sub-tests and timed separately, the questions and problems in AGCT were incorporated in one test form and administered with a single time limit. Instructions to examinees were clear and specific, and a fore-exercise acquainted those unfamiliar with tests of this sort with what to expect. Selection and scaling of test items was carried out on a much more extensive scale than was done with Alpha. AGCT was designed to measure native capacity, but like Alpha and Beta it was influenced by schooling, socio-economic status, and cultural opportunities plus degree of coöperation, attitude toward the test, and other temperamental factors.

The materials in AGCT had frequently been used in well-known intelligence tests, so that validation of this examination as a measure of intellect was not as necessary as it was in the case of Army Alpha. The correlations of AGCT with other measures of intelligence were calculated, however, and proved to be quite high: .73 with school grade completed, .90 with Army Alpha, .83 with the Otis Higher Mental Ability Examination, .30 to .40 with the grades achieved in a number of training programs. AGCT is not related to chronological age, and hence is useful over a wide age range.

Like Alpha, the reliability of AGCT was quite high, the correlation of two forms of the test being around .95. The chances are even, therefore, that upon repeating the test a soldier's score will not shift by more than five points up or down.

Intelligence examinations much like AGCT were widely used in the Navy and Air Force during World War II. In addition, tests of mechanical aptitude, motor and manual dexterity, speed and coordination, and visual acuity, were constructed for use in the selection of pilots, bombardiers, radio operators, clerks, teamsters, cooks, and men in other specialized jobs. Special aptitude tests of this sort were virtually unknown in World War I. They provided a valuable supplement to the AGCT in the selection of personnel for special duties and for job training.

2. CONSTRUCTION OF A GROUP INTELLIGENCE TEST

Army Alpha consisted of eight tests: (1) following directions; (2) arithmetic problems; (3) practical judgment; (4) synonym-antonym; (5) disarranged sentences; (6) number-series completion; (7) analogies; (8) general information. The items in each of these sub-tests were arranged to be progressively more difficult from the beginning to the end of the test. Time limits were set for each test short enough to prevent any but the fastest workers from finishing. This precaution is necessary as the man who finishes the test before time is called might have done more had the material been available. Contrariwise, the man who scores zero on the test might have done a few items at least, had still easier ones been provided. Five forms of Alpha were constructed, all approximately equal in difficulty.

What Alpha was designed to do will be clearer if we examine sample items from the eight sub-tests.

(1) In Test 1, following directions, each item was to be marked by the soldier in accordance with directions given by the examiner. The following is an example.

Attention! Look at 2, where the circles have numbers in them. When I say "Go," draw a line from Circle 1 to Circle 4 that will pass *above* Circle 2 and *below* Circle 3. Go! (Allow not over five seconds.)

(2) Test 2, arithmetic problems, consisted of twenty ordinary problems in mental arithmetic. The tenth problem read as follows:

10. If it takes 6 men three days to dig a 180-foot drain, how many will dig it in half a day? Answer ()

Five minutes were allowed for this test.

(3) Test 3, practical judgment, was intended to be a measure of common sense. Directions were to indicate the best of three possible answers to a given question by placing an "X" in the small box before it. Item 7 proposes the following:

7. Why is wheat better for food than corn? Because
☐ it is more nutritious
☐ it is more expensive
☐ it can be ground finer

One and one-half minutes were allowed for the sixteen items in this test.

(4) Test 4, synonym-antonym, was designed to gauge the ability to grasp relations of likeness and difference. Comprehension of abstract relations is believed by many psychologists to be a highly important aspect of intelligent activity. Items 26 and 36 are examples:

26. Fallacy-verity same—opposite
 36. Innuendo-insinuation same—opposite

Instructions are to underline "same" if the two words mean the same or nearly the same; otherwise to underline "opposite." There were forty items in this test, and one and one-half minutes were allowed.

(5) Test 5, disarranged sentences, measured ingenuity and cleverness as exhibited in the rearrangement of jumbled words into a sentence. Items 16 and 21 are illustrations:

16. ninety canal ago built Panama years was the true—false
 21. employ debaters irony never true—false

The twenty-four items in this test were answered by drawing a line under "true" or "false." The time allowance was two minutes.

(6) Test 6, number series completion, was designed to measure "reasoning ability." The task was to complete a series of numbers in a logical fashion. The thirteenth and sixteenth items read:

- | | | | | | | | |
|----|----|----|----|----|---------------|------|------|
| 11 | 13 | 12 | 14 | 13 | 15 | | |
| 81 | 27 | 9 | 3 | 1 | $\frac{1}{3}$ | | |

Directions are to write on the dotted lines the two numbers which should come next in the series. There were twenty items in the test, and three minutes were allowed.

(7) Test 7, analogies, is another test of the ability to reason or see relations expressed in a verbal form. To illustrate with items 17 and 36:

17. lion—animal :: rose—smell leaf plant thorn

36. tolerate—pain :: welcome—pleasure unwelcome friends give

The examinee must find the relationship between the second and the first words and then underline one of the four alternatives which is related in the same way to the third word. There were forty items, and a time allowance of three minutes was given.

(8) Test 8, general information, was included in order to provide an estimate of the extent to which the individual has profited from contact with his surroundings. This test has been criticized on the grounds that it draws upon superficial and casual observation and is not a test of intellectual ability. It proved, nevertheless, to be a good test for separating officers from enlisted men. Given a common adult environment, it seems reasonable to suppose that the mentally alert will pick up more information than will the dull and stupid. Items 21 and 37 will serve as illustrations:

21. The dictaphone is a kind of typewriter multigraph phonograph adding machine

37. Mauve is the name of a drink color fabric food

Instructions are to underline that one of the four possible answers which makes the best or "truest" sentence. There were forty items in all, and four minutes were allowed.

A revision of Army Alpha was published in 1941. Unsatisfactory, ambiguous and outdated items were eliminated and the sub-tests so arranged that separate verbal and numerical scores can be calculated.

The Army Beta Test consisted of seven non-language sub-tests, and was intended for use (p. 245) with illiterates and non-English-speaking soldiers. The tests in Beta may be described as follows: (1) maze drawing; (2) cube analysis; (3) X-O series, in which the examinee completed a series of X's and O's arranged in different sequences; (4) digit-symbol; (5) number checking; (6) picture completion; (7) geometric construction. Each of these tests was timed separately, approximately thirty minutes being allowed for the whole test. The examiner explained each test by gesture and pantomime, and by use of a large demonstration screen.

The AGCT was set up in four equivalent forms. There are no sub-tests, the items appearing in regular sequence. Including instructions and fore-

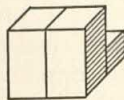
exercise, the overall testing time is about one hour. AGCT attempts to measure three important aspects of intellect: verbal comprehension, quantitative reasoning, and spatial perception. Proficiency in each of these abilities has been found to be highly important in learning and in the solution of abstract and technical problems. Three kinds of test items, vocabulary, arithmetic problems, and block counting, were employed to measure achievement in the three specified areas. Questions were arranged in spiral fashion, first vocabulary, then arithmetic, then blocks, say. There were 150 items in the test as a whole (fifty of each kind), arranged in order of difficulty; all were of the multiple-choice type. The following samples from the practice booklet will serve as illustrations:

To PERMIT is to (A) demand (B) thank (C) allow (D) charge
Tom sold 18 pints of milk at 9 cents a quart. How much money did he get for the milk?

- (A) 50 cents (B) 81 cents (C) \$1.00 (D) \$1.62

How many blocks?

- (A) 5 (B) 4 (C) 3 (D) 6



After World War I, a number of group tests, modeled after Alpha, were constructed for use with children and adults. Today we find intelligence tests designed for pupils in elementary school, for students in high school and college, and for adults beyond school age. Examples of widely used examinations are: The Pintner-Cunningham Primary Test (kindergarten—grade 2); the Kuhlmann-Anderson Intelligence Tests (grades 1–12); the Terman-McNemar Test of Mental Ability (grades 7–12); the American Council on Education Psychological Examination (for college freshmen and high school students).

Intelligence tests differ considerably in content and form. Commonly found sub-tests are arithmetic problems, opposites, analogies, vocabulary, block counting, picture arrangement, and classification. Items in a group intelligence test are sometimes organized into sub-tests and are sometimes printed in a single series or list following a standard or cyclic order—so that an arithmetic problem, say, recurs at regular intervals in the test. Norms or performance standards, together with descriptions, directions for administering, and the like, will be found in the manual accompanying the test.

3. THE INTELLIGENCE OF THE SOLDIER

The score on Army Alpha was obtained by adding together separate scores on the eight sub-tests. The maximum possible score was 212. When

a soldier's total score had been found, it was translated into a letter grade as shown in Table XXV.

TABLE XXV
SCORE RANGES AND LETTER GRADE EQUIVALENTS ON THE ARMY
ALPHA EXAMINATION

Letter rating	A	B	C+	C	C—	D	D— and E
Range of Alpha scores corresponding	212—135	134—105	104—75	74—45	44—25	24—15	14—0
Per cent of white soldiers receiving each letter grade	4	8	15	25	24	17	7

Letter grades corresponding to the different score ranges were found by drawing up a distribution of Alpha scores made by a large and representative group of soldiers. This distribution was then subdivided into five ability sections: (A) high, (B) fairly high, (C+ and C) average, (C—) low, and (D, D— and E) very low. Assigned letter grades are measures of relative performance and should not be confused with absolute measures such as inches and pounds. Note that any score from 212 to 135 receives an A rating; any score from 134 to 105 a B rating; and so on. About 4 per cent of the white soldiers were given A ratings, and about 12 per cent A and B ratings, while nearly two thirds fell into the C+, C and C— groups. The mean score made by white enlisted men was 59 Alpha points, which corresponds to a letter rating of C. White officers made on the average a score of 139, nearly all of them earning letter ratings A or B.

Shortly after World War I, the unqualified statement was made by several popular writers that the typical American soldier had the mentality of a fourteen-year-old child. To many people this was a matter of grave concern. They regarded with uneasiness and even dismay the precarious outlook for the nation in view of the mental immaturity of a large share of its citizens. Fortunately, this alarming view was entirely unwarranted: it grew out of a misunderstanding of what the Army tests do and of what "mental age" means. Let us look at the method whereby the average mental age of 14 years was obtained. Through the correlation of Alpha and Stanford-Binet it was possible to convert Alpha scores into Stanford-Binet mental ages. When this was done the average score of 59 made by enlisted men on Alpha became equivalent to a Stanford-Binet mental age of approximately 14. However, MA (like IQ) is a measure of mental development—meaningful when used with children but dubious and misleading when applied to adults. A feeble-minded man of forty, for example, behaves very differently from a normal child of ten, even though both have Stanford-Binet mental ages of 10.

The statement of Alpha score in terms of mental age, then, means simply that the average soldier did about as well on the test as the average first-year high school child. Certainly there is nothing disturbing about this finding. Reference to Figure 39, p. 226, shows that the rise in the age-progress curve for normal intelligence from 14 to 16 is relatively small. The typical soldier was about two years below the MA level of 16 set by the 1916 Stanford Revision as the age at which intelligence matures in the average person.¹ It must be remembered, however, that the schooling of many of these inductees had ended years before, which put them at a disadvantage when compared with school children on a written test. Moreover, specialized training and the routine of adult occupation often play havoc with whatever skill a man may have had at one time in reading and figuring rapidly.

It may be added further that Alpha gave no measure of manual skill, of ability to exercise good judgment, to control one's temper, or to get along on friendly terms with one's associates. Mental alertness as shown in the ability to solve problems involving language and numbers is undeniably important. But skill in dealing with ideas does not run the gamut of achievement. Hence we need not be greatly perturbed that the average soldier in World War I earned a mental age of 14 years on a scale intended for school children.

In World War II, Army psychologists did not make the mistake of scoring the AGCT in either mental age or IQ units. Instead, point scores earned on the test (which might range from 0 to 150) were converted into a standard scale in which the mean was set at 100 and the SD 20.² On the basis of his standard score, a soldier was classified into one of five Army grades, in accordance with the following table.

TABLE XXVI
SHOWING SCORE RANGES IN AGCT FOR, AND PER CENT OF, SOLDIERS
IN EACH ARMY GRADE

<i>Army grade</i>	<i>AGCT standard score range</i>	<i>Per cent expected in Army grade</i>
I	130 and above	7
II	110-129	24
III	90-109	38
IV	70-89	24
V	Below 70	7

The figures in the last column of Table XXVI give the expected frequencies of men in each Army grade. Actually, the proportion of men

¹ In the 1937 Stanford Revision, the point of maturity is 15 years (p. 224).

² The SD is a measure of the variability or range in a distribution of scores. The bulk of AGCT scores fell between 40 and 160, that is, within ± 3 SD from the mean of 100.

tended to be somewhat higher in grades IV and V and often lower in grades I, II, and III (see p. 259).

The distribution of AGCT standard scores for 9,339,289 soldiers is shown in Figure 43. While the curve is quite symmetrical around the mean of 100, there is a piling up of scores at the low end of the distribution (below 60). In order to do anything at all with the AGCT a man should have at least a fourth-grade education. Inclusion in the test group of illiterates and those both feeble-minded and illiterate leads to an accumulation of low scores. Low scoring men in AGCT were generally reexamined with non-language tests, just as the men who scored low on Alpha in World War I were later given Beta.

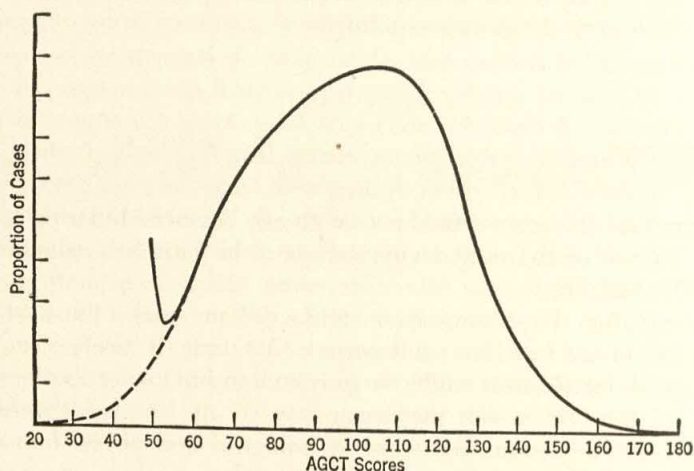


FIG. 43
DISTRIBUTION OF SCORES MADE BY 9,339,289 MEN ON AGCT
(J. ed. Psychol. 1947, 38, p. 414.)

An interesting comparison of the performance of soldiers in World Wars I and II can be made from the Alpha scores achieved by a sample of World War II inductees. The average soldier of World War II earned a much higher score on Alpha than did his World War I counterpart. In fact, only 17 per cent of the first War soldiers did as well on Alpha as the upper half of second War soldiers. This better performance after twenty-five years can be attributed in part at least to longer and better schooling. The average educational level of World War II soldiers was two years of high school as compared with an eighth-grade average for World War I soldiers. Also, in the 1940's young men had the advantage of better nutrition, better health, and more experience with objective tests than soldiers in 1917. English was spoken with greater facility, and there was less foreign language interference to lower test scores.

4. INTELLIGENCE AND AGE

Results from the Army tests throw light upon a problem important from both the social and economic point of view, namely, the question of whether intelligence decreases with age. We have seen, on page 226, that according to the best available evidence mental growth in the normal child reaches its peak at 15 to 16 years. Since the individual intelligence test is intended primarily for use with children, group tests are valuable in showing changes in intelligence with age among adults. From the late 'teens well on into maturity, general intelligence as measured by group tests shows a gradual but steady decline. This may be seen in the table below which gives by decades the median ³ scores of Army officers upon Alpha.

TABLE XXVII

Age-groups	21-30	31-40	41-50	51-60
Median score on Army Alpha	145	133	125	120

The gradual decrease of test score with age has been reported in many studies. In one of the most comprehensive, the Otis Self-Administering Test of Mental Ability was administered to 823 persons (both men and women) ranging in age from seven to ninety-four years. Figure 44 shows the changes in test score over this period. Note that the decline is gradual from 20 to about 40, after which there is a more pronounced falling off.

Some of the loss in test performance with age can undoubtedly be attributed to the slowing down of physiological and physical functions, but this is probably not the only—or even the most important—cause. Older folk, to be sure, are likely to have poorer health and more visual and auditory defects than younger. But older adults show greater variation in educational and vocational background than younger, and tend to be more specialized in their knowledge and interests. Moreover, they are further removed from formal schooling and from practice in the rapid performance of tasks; and are likely to be less coöperative and less interested in taking tests under timed conditions.

It is significant that the loss in mental ability with age as indicated by test score is relatively slight in those measures which depend upon previous knowledge and information, and is relatively large in tasks demanding quickness and versatility. When the Alpha sub-tests are studied separately, those depending upon alertness show a greater loss with age than those requiring reasoning and background knowledge. Loss in score with age is decidedly variable, too, many older persons being far more

³ The median or mid-score is the middlemost when scores are arranged in order from lowest to highest. Thus it is typical of the group.

able than those who are much younger. Statistics on the intellectual productivity of eminent men show that the greatest output of authors, inventors, and scientists is between the ages of 25 and 40. Masterpieces are most often produced between 40 and 50 years, but many eminent men have done brilliant work long past this time. Bismarck was in office at 75; Gladstone was Prime Minister of Great Britain at 85; Edison was taking out patents on inventions at 80; Titian painted some of his best work after 70; and there are many others.

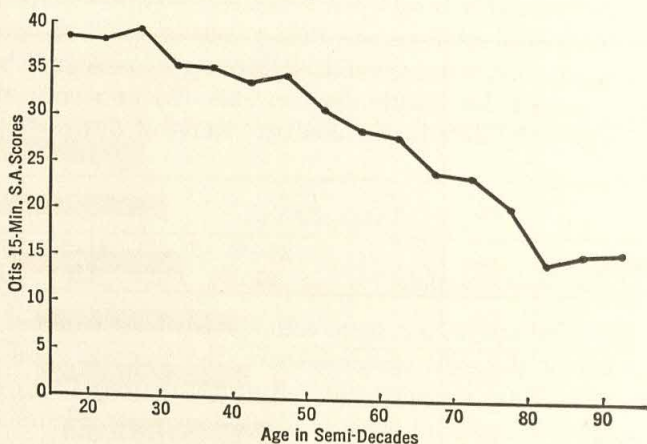


FIG. 44

DECLINE IN INTELLIGENCE TEST SCORES AFTER MATURITY
(From Miles & Miles, 1932.)

Another fact worthy of comment is that the higher the educational level attained in youth the smaller the loss in mental test score with age. This probably means that the higher the initial level of mental ability the more slowly does it decline with the years. Furthermore, the well educated man has more interests than the uneducated, surrounds himself with a more stimulating environment, and develops more resources within himself. All of these things serve to maintain enthusiasm and prevent mental stagnation as the physical activities wane.

5. INTELLIGENCE AND OCCUPATIONAL LEVEL

Men inducted into the Army in World War I were required to list their occupations in civil life, and these army records, together with Alpha scores, provided a basis for the first comparative study of the intelligence of men in different occupations. As might be expected, we find an occupational order or hierarchy, the highest scores being made by professional groups, the lowest by unskilled laborers. In the groups scoring A

and B on Alpha are civil engineers, physicians, lawyers, teachers, and business men. In the C+ group are stenographers, bookkeepers, clerks, photographers, and workers at skilled trades. In the C group we find carpenters, policemen, tailors, butchers, printers, farmers, and small storekeepers. Store clerks, cooks, fishermen, firemen, barbers, and day laborers made up the lowest group (C- and D).

Analysis of AGCT scores when classified by occupation agrees closely with the occupational orders given earlier by Alpha. Figure 45 shows the mean AGCT score and the ranges of score for fifteen selective occupations.

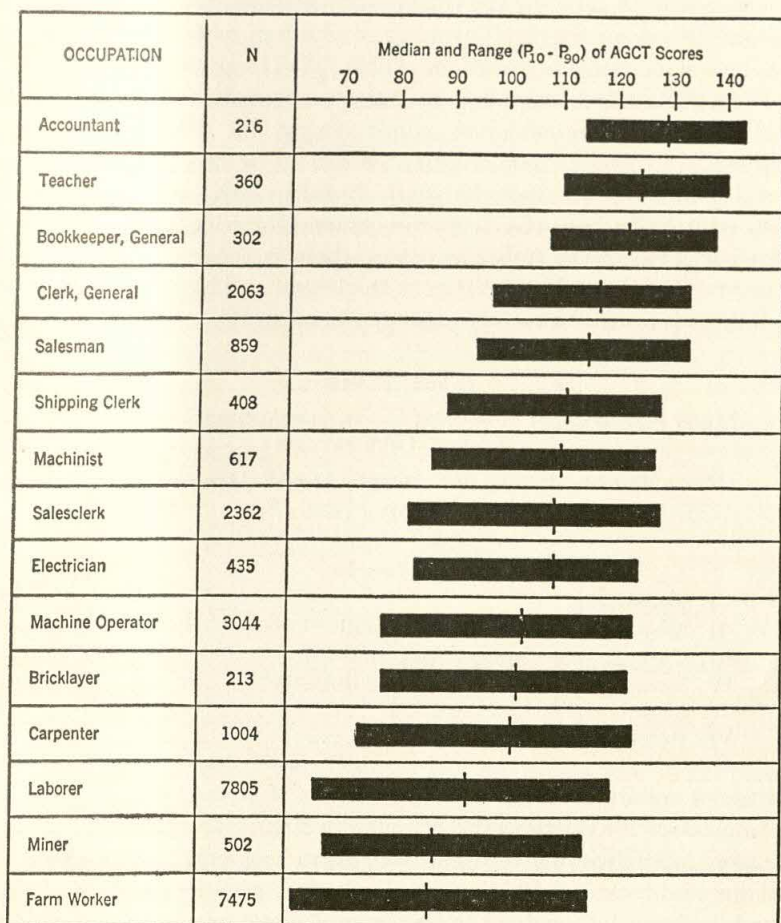


FIG. 45

AGCT SCORE IN RELATION TO CIVILIAN OCCUPATION

(From Anastasi and Foley, *Differential Psychology*, rev. ed., 1949. The Macmillan Company.)

The profiles in the Figure were based upon the scores made by a 2-per cent sample drawn at random from the entire Army personnel. The bar graphs in Figure 45 are not strictly representative of the male American civilian population. The age range is restricted and numbers of men have been deferred, rejected or discharged for various reasons. Moreover, no data from professional groups (officers, physicians, and the like) are included. It is interesting to note that in spite of these restrictions an occupational hierarchy clearly appears. Intelligence test score parallels closely the socioeconomic level of the occupation.

Another way of showing the relationship between intelligence and occupation is to compare the IQ ratings of children whose fathers are engaged in a wide range of vocations. Table XXVIII gives the average IQ's of children classified according to father's occupation. These data are based upon results obtained by administering the 1937 Stanford-Binet individually to nearly 3,000 children chosen to be representative of the general population of American youth. Results closely parallel those obtained with Alpha and AGCT. As the demand for education and for specialized training of the father increases, there is a corresponding rise in the mean IQ of the children. Between the lowest and highest occupational levels there is roughly a twenty-point gap in mean IQ.

TABLE XXVIII

MEAN IQ'S OF 2,757 CHILDREN CLASSIFIED ACCORDING TO THEIR FATHERS' OCCUPATIONS

(From Terman, L. M., and Merrill, M. A., *Measuring Intelligence*, 1937, pp. 14 and 48.)

<i>Fathers' Occupational Level</i>	<i>Children's Mean IQ</i>
I. Professional	116.2
II. Semi-professional and managerial	111.9
III. Clerical, skilled trades, retail business	107.5
IV. Semi-skilled, minor clerical and business	105.0
V. Slightly skilled and unskilled	97.2
VI. Rural owners	95.1

It seems certain that native endowment as well as training must operate to produce the relationships revealed in Figure 45 and Table XXVIII. The more intelligent men seek and are able to hold positions in which professional and technical training are demanded, and the children of these more intelligent fathers tend to be above average in mental test ratings.⁴ Moreover, these children have a higher social and economic status, in gen-

⁴ The correlation (p. 269) between intelligence test scores in parent and offspring is about +.50.

eral are better educated, and are exposed to a more cultured environment. By the same token, the less intelligent men enter occupations requiring less training and less educational equipment *because* they are less able intellectually. In general their homes lack the intellectual stimulation characteristic of the better educated groups; and this deprivation plus poorer heredity is reflected in the scores of the children.

Everyday observation confirms the wide differences in mental level found among occupational groups in Alpha and AGCT. It must not be forgotten, however, that the range of intelligence within any occupational level is wide, and that men of high intelligence are found in occupations for which the average rating is low. Poor environment, lack of formal education, opportunity and temperament, as well as other less tangible influences, are important in effecting this result. Environmental factors must always be considered in judging a man's probable intelligence from his occupation; they qualify but do not discount the importance of the relationship between occupation and intelligence disclosed by the Army tests.

We have seen that officers on the average achieve higher Alpha scores than enlisted men. This result was to be expected as was also the finding that officers in branches of the service which require technical training and special preparation rank highest in the intelligence test. Figure 46 shows in striking fashion the chances that a soldier who achieves a given grade on AGCT will receive a commission in officer candidate school. An AGCT score of 110 was decided upon as the minimum requirement for entrance to officer training courses. It is clear from Figure 46 that men scoring 130 and above have a much better chance of being commissioned than men scoring under 110. In fact, a man scoring 140+ is twice as likely to earn a commission as is the man who scores under 110.

6. NEGRO-WHITE DIFFERENCES ON THE ARMY TESTS

Owing to the size of the groups and lack of special selection, Alpha and Beta yielded the fairest and most unbiased comparison of Negro and white intelligence in the United States up to 1917-1918. Table XXIX shows that Negro soldiers scored lower than whites on Alpha, southern Negroes lower than northern Negroes. This is true not only for Alpha (a language test) but also for Beta (non-language). Scores on the two examinations cannot be compared directly as the tests are scored in different units, and five points on one does not mean the same as five points on the other. A comparison of the score distributions of the two groups, however, reveals that approximately 25 per cent of northern Negroes exceeded the median white, whether the test was Alpha or Beta; and that 5 per cent of southern Negroes exceeded the white median on Alpha and 10 per cent the white

median on Beta. It is clear that the Negro soldier lagged behind the white to about the same extent in both tests. Owing to the large overlap in test score in the two racial groups, however, many Negroes will be found who are more intelligent than the average white. This is especially true where the selection of Negroes is stringent as in the case of officer material. For instance, while only 40 per cent of Negro officers in World War I achieved A or B ratings on Alpha as against 80 per cent of white officers, only 12 per cent of white enlisted men achieved A or B ratings.

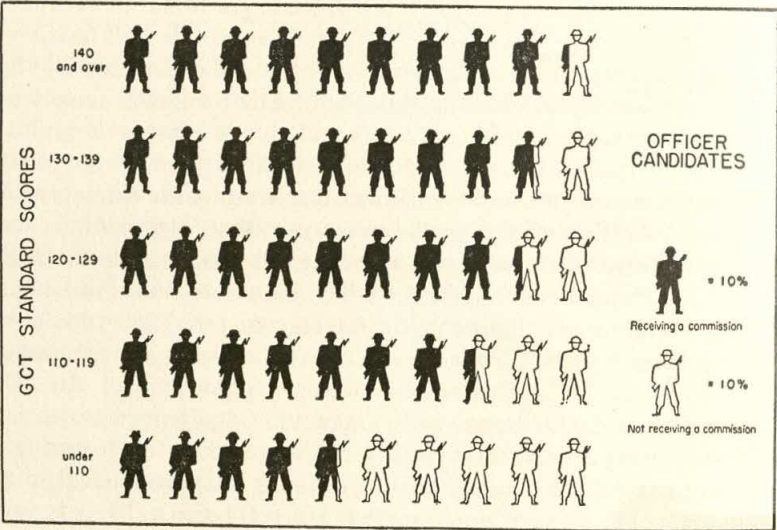


FIG. 46

PREDICTING SUCCESS OF OFFICER CANDIDATES FROM THEIR GRADES IN THE ARMY
GENERAL CLASSIFICATION TEST
(From *Psychology for the Armed Services*, 1945, Infantry Journal.)

In World War II, from 15 to 20 per cent of Negro soldiers exceeded the white median in AGCT, the comparative performance of the two groups being about what it was twenty-five years earlier. Of the inductees in the nine service commands who were processed from June 1943 to May 1945, 76 per cent of the Negroes fell in Army groups IV and V, as compared with 30 per cent whites in the same grades; 7 per cent of the white

TABLE XXIX

MEDIAN SCORES ON ALPHA AND BETA OF NEGRO AND WHITE SOLDIERS

	Alpha	Beta
White	59	43
Northern Negro	39	32
Southern Negro	12	20

soldiers as compared to .4 of 1 per cent Negroes were classified in Army group I. (See p. 251 for theoretical expectation.)

Various reasons have been advanced in support of the view that Negro-white differences in intelligence test scores are environmental and offer no evidence of innate or genetic differences. It has been pointed out, for example, that verbal intelligence tests are known to be highly correlated with educational and cultural opportunities, in both of which Negroes are generally below whites; that Negroes achieve scores at all levels of the intelligence range; that regional differences are often greater than racial differences; and that as educational and cultural environment are more nearly equalized, the performance of the Negro approximates more closely that of the white.

TABLE XXX

PERCENTAGES OF NEGRO AND WHITE SOLDIERS IN EACH ARMY GRADE (AGCT)
FOR DIFFERENT LEVELS OF EDUCATION (After Davenport)

(Draftees processed from June 1941 to February 1942.)

N = 549,608

<i>Schooling</i>	<i>Race</i>	<i>I + II</i>	<i>III</i>	<i>IV + V</i>
Some schooling	White	4%	18.9%	77.3%
	Negro	1.3	4.9	93.1
Completed Grade school	White	23.8	42.8	33.4
	Negro	3.9	19.4	76.5
Completed High school	White	65.6	28.5	5.9
	Negro	18.1	41.5	40.3
Completed College	White	92.5	7.1	.4
	Negro	41.3	44.3	14.4
College Post-graduate	White	93.5	6.0	.5
	Negro	68.8	25.0	6.3

While the case for environment is undoubtedly strong, there is evidence that it is not the whole story. Table XXX gives the percentage of Negro and white inductees at different levels of education who fell in Army grades I + II, in III, and in IV + V over the period from June 1941 to February 1942. Approximately 550,000 men are represented in this sample. It will be noted that the proportion of whites in army grades I + II exceeds by a considerable margin the proportion of Negroes—and this is true when schooling has been presumably equated for the two groups. Moreover, the proportions of Negroes in Army grades IV + V are much larger than the proportions of whites in these same grades—for each level of education. Even the whites who completed grade school do better than the Negroes who completed high school.

One point should be stressed in conclusion. The mental test differences cited are *average* differences which apply to Negro-white groups as we find them in America, not to groups in which *all* educational, cultural, and social differences have been eliminated. That the Negro would test lower than the white when *all* such factors are equated we do not know; that he does test lower under the conditions in which we find him is a demonstrated fact.

7. NATIONAL DIFFERENCES IN INTELLIGENCE

An interesting comparison of the intelligence ratings of various nationalities may be obtained from the Alpha scores achieved by foreign-born men inducted into the American Army in 1917-1918. The scores of various national groups were expressed in terms of a "combined scale"—made up of the eight Alpha tests, Stanford-Binet, and four tests from Beta. The maximum score on the combined scale was 25 points. Table XXXI gives the mean scores of men born in various foreign countries together with the number of men in each group.

TABLE XXXI

<i>Country of Birth</i>	<i>Number of Men</i>	<i>Mean Intelligence Score</i>
England	411	14.87
Scotland	146	14.34
Holland	140	14.32
Germany	308	13.88
Denmark	325	13.69
Canada	972	13.66
Sweden	691	13.30
Norway	611	12.98
Belgium	129	12.79
Ireland	658	12.32
Austria	301	12.27
Turkey	423	12.02
Greece	572	11.90
Russia	2,340	11.34
Italy	4,009	11.01
Poland	382	10.74

It is apparent from the table that men born in northern European countries ranked higher than men born in southern European countries, though the differences are small. Publication of these scores led to acrimonious discussion after World War I. Supporters of the theory of Nordic superiority held that Nordics (north Europeans) are more intelligent than Alpines (mid-Europeans) who in turn excel Mediterraneans (south Euro-

peans): and they pointed to the results of the intelligence tests as experimental evidence for this claim.

Opponents of the Nordic theory attributed variations in intelligence test score to differences in language, customs, training, and educational background. They argue that we do not know how representative of each country these small samples were, as American immigration has rarely been from the same social and intellectual strata of the various European countries.

8. INTELLIGENCE AND SCHOOLING

A soldier's intelligence test score is closely related to his amount of schooling. The correlation between Alpha and school grade completed was .74 in a sample of approximately 50,000 men, and the correlation of AGCT and school grade was .73 in a group of 4,330 inductees.

Table XXXII gives the median Alpha scores made by officers and enlisted men in World War I who had completed four or less school grades, five to eight school grades, high school and college.

TABLE XXXII

*Median Alpha Scores for
Given Amounts of Education*

<i>Group</i>	<i>0 to 4 grades</i>	<i>5 to 8 grades</i>	<i>High school</i>	<i>College</i>
White officers	112.5	107.0	131.1	143.2
White soldiers (native-born) ..	22.0	51.1	92.1	117.8
White soldiers (foreign-born) ..	21.4	47.2	72.4	91.9
Negro soldiers (Northern)	17.0	37.2	71.2	90.5
Negro soldiers (Southern)	7.2	16.3	45.7	63.8

The importance of educational level to success in Alpha is shown by the rise in Alpha score in all groups (with a single exception) as amount of schooling increases. In fact the general rule is, the more the schooling the better the intelligence score. Reading down the columns, however, we find striking evidence of the influence of native ability. Officers scored consistently higher than native born white soldiers for each grade of education, while the latter group in turn regularly exceeded the other three groups. These regular and consistent differences in Alpha score for a constant amount of schooling bear significant testimony to the part played by native ability.

The influence of schooling upon AGCT score appears clearly in Table XXX. Score rises steadily with increase in educational level, and this is true for Negroes as well as for whites. The fact that the more the schooling the

higher the test scores, is doubtlessly the result to be expected, as education is in itself a highly selective process. A study of the school records of 80,000 inductees in World War I showed that of every 1,000 who entered the first grade, 830 reached grade five, 490 grade eight, 95 graduated from high school, and 10 graduated from college. By contrast, in 1947 (World Almanac, 1949), 220 out of every 1,000 persons twenty through twenty-four years old had completed the grades, 650 had finished high school, and 130 had taken from one to four years of college work. This is a vast improvement over the record thirty years earlier. There are, to be sure, many reasons why students fail and drop out of school, but certainly one of the most potent is sheer inability to do school work satisfactorily with resulting discouragement and loss of interest.

9. INTELLIGENCE TESTS IN THE COLLEGES

Immediately after World War I, Alpha was given to hundreds of students in our colleges and universities. The median Alpha score made by college freshmen fell between 130 and 140. This compares favorably with the median Alpha score of 139 made by white officers and is enormously higher than the 59 made by the average soldier.

Alpha is not a good test for college students. Planned to measure intelligence over a wide range of performance, Alpha does not measure high enough, is generally too easy, and tends among the better students to become little more than a speed test. The AGCT is useful when range of ability is wide, and there are many good tests designed specifically for the college level. For example, the American Council on Education Psychological Examination, the Ohio State Psychological Test, the Yale Educational Aptitude Battery, and the tests of the College Entrance Examination Board are widely used and superior instruments for the selection and placement of college students.

BIBLIOGRAPHY

1. Boring, E. G., Ed. Psychology for the armed services, National Research Council. Washington, D.C.: Infantry Journal, 1945.
2. Brigham, C. C. A study of American intelligence. Princeton: Princeton University Press, 1923.
3. Bruce, M. Factors affecting intelligence test performance of whites and Negroes. Arch. Psychol., New York, 1940, 36, No. 252.
4. Davenport, R. K. Implications of military selection and classification in relation to universal military training. J. Negro Educ., 1946, 15, 585-594.
5. Franzblau, R. N. Race differences in mental and physical traits. Arch. Psychol., New York, 1935, 26, No. 177.
6. Hirsch, N. D. M. A study of natio-racial mental differences. Genet. Psychol. Monog., 1926, 1, 231-406.

7. Intelligence: its nature and nurture. Thirty-ninth Yearbook of the National Society for the Study of Education, 1940, Parts I & II.
8. Jones, H. E., & Conrad, H. S. The growth and decline of intelligence: a study of a homogeneous group between the ages of ten and sixty. *Genet. Psychol. Monog.*, 1933, 13, 223-298.
9. Klineberg, O. A study of psychological differences between "racial" and national groups in Europe. *Arch. Psychol.*, New York, 1931, 20, No. 132.
10. Klineberg, O. Negro intelligence and selective migration. New York: Columbia University Press, 1935.
11. Miles, C. C. & W. R. The correlation of intelligence scores and chronological age from early to late maturity. *Amer. J. Psychol.*, 1932, 44, 44-78.
12. Miles, W. R. Age and human ability. *Psychol. Rev.*, 1933, 40, 99-123.
13. Noble, E. L., & Arps, G. F. University students' intelligence ratings according to the Army Alpha test. *Sch. & Soc.*, 1920, 11, 233-237.
14. Staff, Personnel Research Section, Classification and Replacement Branch, A.G.O., Personnel Classification Tests, War Department Technical Manual TM, 12-260 rev. U.S. Printing Office, 1946.
15. Staff, Personnel Research Section, A.G.O. The Army General Classification Test, with special reference to the construction and standardization of forms 1a and 1b, *J. Educ. Psychol.*, 1947, 38, 385-420.
16. Stewart, N. AGCT scores of army personnel grouped by occupation, Occupations, 1947, 26, 5-41.
17. Terman, L. M., & Merrill, M. A. Measuring intelligence. Boston: Houghton Mifflin Company, 1937.
18. Tuddenham, R. D. Soldier intelligence in World Wars I and II. *Amer. Psychol.*, 1948, 3, 54-56.
19. Tyler, L. E. The psychology of human differences. New York: Appleton-Century-Crofts, Inc., 1947.
20. Yerkes, R. M. (Ed.) Psychological examining in the United States army. *Memoirs National Academy of Sciences*, 1921, XV. Washington: Government Printing Office. Pp. 162, 422, 764, 766, 767, 768, 770, 779, 800, 814, 824, 836, 848.
21. Yoakum, C. S., & Yerkes, R. M. Army mental tests. New York: Henry Holt and Company, 1920.

GALTON AND THE MEASUREMENT OF INDIVIDUAL DIFFERENCES

1. THE BEGINNING OF INTEREST IN INDIVIDUAL DIFFERENCES

THE EMINENT English scientist, Francis Galton (1822–1911) is probably best known to psychologists as a pioneer student of heredity and as the first investigator to apply statistical methods systematically to the problem of individual and group differences in mental traits. Galton was not a psychologist in the professional sense. He did not hold an academic position, nor did he confine himself to work upon problems primarily psychological. His interests were wide and varied. According to Terman, Galton showed the marks of genius in early childhood and his later versatility, originality, and active interest in problems of all sorts bear out this early promise. Galton was the founder of the eugenics movement; he established one of the very first laboratories (in 1884) wherein mental and physical tests—mostly of the sensory-motor sort—could be taken for a small fee; he invented or devised various instruments for measuring sensory acuity (e.g., the Galton Whistle); he introduced the rating scale and questionnaire methods later so widely used; and he initiated and fostered the statistical study of individual differences.

Galton's work influenced and was influenced by psychological research in other countries. In 1894 the German psychiatrist Kraepelin established a psychological laboratory for the examination of mental patients in connection with his clinic and hospital. Instead of a vague and general description of symptoms, Kraepelin insisted upon precise measurement, the use of objective "tests" and the importance of measuring deviations from normal behavior. Kraepelin was interested in the magnitude, the range, and the meaning of individual differences. Reaction time, the speed and character of verbal association, memory, perception, sensory acuity, discrimination, and fatigue were among the activities and performances measured in his laboratory. William Stern was another eminent German pioneer in the field of individual differences. Stern's book on differential psychology, first published in 1900, has since appeared in several editions. Stern dealt with

differences among people in various psychological traits, from simple sensory capacities to the more complex mental and emotional characteristics. The differences found among various racial and cultural groups as well as between occupational and social levels were analyzed and evaluated.

Under Binet's influence, differential psychology in France was mainly concerned with the measurement of complex activities and in the development of intelligence tests (Chap. 11). Binet's approach differed from the German and English in its objectives and in the kinds of tests used in measuring psychological functions. Both movements, however, grew out of a common interest in the likenesses and differences exhibited by individuals, and in the reasons for these variations.

Mental testing in the United States was a direct outgrowth of the European work upon individual differences. One of the leaders in this field was J. McKeen Cattell, a student of Galton and a pioneer both in the development of experimental psychology and of mental measurement. Beginning in 1894 and for several years thereafter, Cattell administered a series of mental and physical tests to Columbia College freshmen. The relationships among these test scores proved to be slight, giving experimental support to the then current theory that abilities are essentially specific in nature. Cattell's tests included among others, strength of grip, rate of movement, perception of differences in weights, reaction time to sound, speed of color naming, bisection of a 50-cm. line, and auditory memory span. These tests, which are patterned after those used by Galton, furnish an interesting contrast to Binet's tests given on pages 221-222.

Other pioneer psychologists in America were interested in differential psychology. At the Chicago Exposition in 1893, Joseph Jastrow administered tests—much like those of Cattell—to interested volunteer subjects. Two years earlier (1891), Munsterburg had employed tests of controlled association, judgment and memory in studies of school children. Gilbert as early as 1893 carried out a fairly extensive investigation of sex differences, of the relation of physical and mental traits, and of growth changes in mental ability. In Gilbert's study, the "general mental ability" of some 1,200 school children was estimated by their teachers, and these estimates compared with scores on eight tests of sensory and motor functions, reaction time, and suggestibility. A slight relationship was found between estimated intelligence and memory, and a fairly marked relationship between intelligence and reaction. Steady improvement was noted in the scores on all tests up to puberty.

The field of differential psychology includes a large number of problems dealing with the variations and diversities found among individuals and groups. In a broad sense all differences among people may be thought

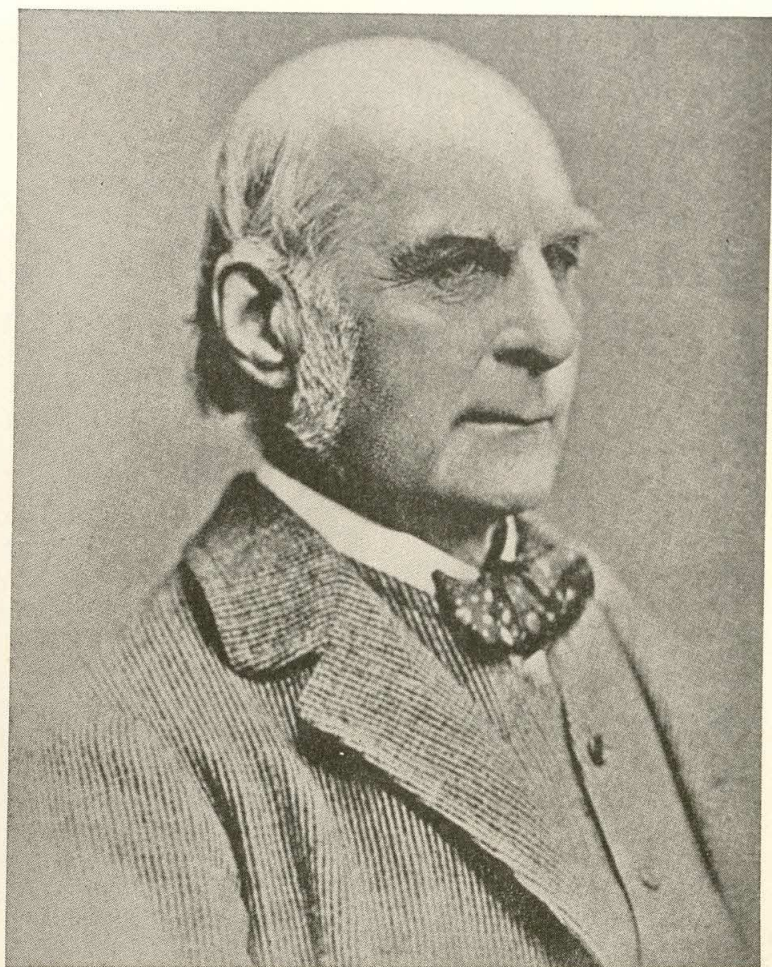
of as joint products of heredity and environment, or, more simply, of nature and nurture. Under nature belong such biological factors as immediate ancestry, race, sex, and age; under nurture all of the social, educational, cultural, and other extrinsic agencies which shape and fashion our behavior from birth until death. It has been a major task of differential psychology to evaluate the relative contributions of these two groups of influences.

2. THE NATURE OF INDIVIDUAL DIFFERENCES

In common parlance, we often speak of human traits as though they occur in "all-or-none" fashion. Thus a pupil is described as bright or dull; a girl as pretty or ugly; a politician as honest or dishonest. Perhaps a third or intermediate category is admitted, but quite frequently only two sharply opposed groupings are recognized. Again, we often speak of a person as the "kind of man" who would—or would not—do this or that. In part, this inveterate tendency to pigeon-hole people with respect to physical and mental traits may be little more than a lazy method of avoiding more precise description. And in part it may be simply a matter of convenience. Definite classification with respect to traits becomes of interest to psychologists when it expresses the belief—admittedly or by implication—in "types," or in sharp natural groupings into which many people fall.

Belief in psychological as well as in body (constitutional) types has had a long history. In an early study of mental imagery, Galton investigated the presence and amount of imagery, and somewhat incidentally the question of "imagery types." Galton used his now famous "breakfast table" questionnaire. Each of his subjects was asked to call up a picture of his breakfast table as it appeared in the morning; and to report whether the objects seemed well defined, natural, and comparable in vividness to the actual objects. Many of Galton's subjects were scholarly men, who to his surprise reported an almost total lack of images. Other subjects reported that the objects on the table appeared in retrospect to be as real as though they were actually present before the eyes. Galton accounts for the feeble imagery of scientific men in the following way: "My own conclusion is, that an over-ready perception of sharp mental pictures is antagonistic to the acquirement of habits of highly-generalised and abstract thought, especially when the steps of reasoning are carried on by words as symbols, and that if the faculty of seeing the pictures was ever possessed by men who think hard, it is very apt to be lost by disuse" (ref. 13).

Interestingly enough, Galton's work offered no evidence of distinct imagery types, though it is sometimes credited with so doing. As a matter of fact, Galton reported that instead of a "visual type," say, he found a



FRANCIS GALTON

1822-1911

gradation in the clearness of visual imagery from vivid to faint. Several investigators who followed Galton took their work to indicate the existence of "imagery types," and widely popularized the notion. The *visile* who is strong in visual imagery but weak or totally lacking in other kinds; the *audile* who is especially strong in auditory imagery; and the *motile*, strong in motor or kinaesthetic imagery, are the types most often cited. Later work, such as that of Betts, who used an expanded form of Galton's questionnaire, cast serious doubt upon the existence of "pure" imagery types. Betts found "mixed" imagery to be the most usual sort, the "pure" visual-minded, auditory-minded, or motor-minded individual being rare if he exists at all.

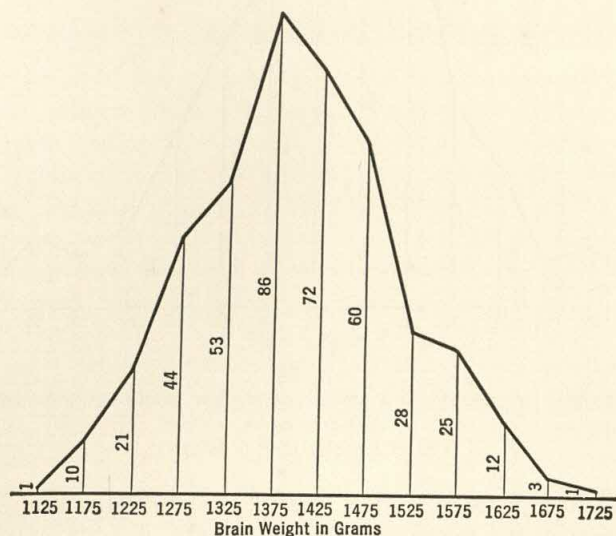


FIG. 47

DISTRIBUTION OF BRAIN WEIGHTS OF 416 MALE ADULTS
(After R. Pearl, 1930, p. 316.)

Discovery that the "mixed" type of imagery is the most common falls in line with other work upon individual differences. Careful measurements of physical and mental traits in large groups of children and adults have given results diametrically opposed to the notion of types. We do not find trait scores occurring in "bunches" or constellations; nor do we find individuals spread evenly over the entire scale. Instead, the usual result is for many individuals to fall at the middle of the scale, and for relatively few to score either very high or very low. Figures 47 and 48 provide typical illustrations of the way in which human measurements are distributed. These two graphs represent the brain weights of 416 adult males, and the vocabulary test scores of 360 eleven-year-old boys.

The two figures are constructed in the same way. To illustrate with Figure 47, brain weights measured in grams and ranging from light to heavy¹ are laid off on the horizontal axis or baseline. Above each weight-point is then marked off, on successive vertical lines, the *number* of individuals having that particular brain weight. A graph or "curve" has been drawn

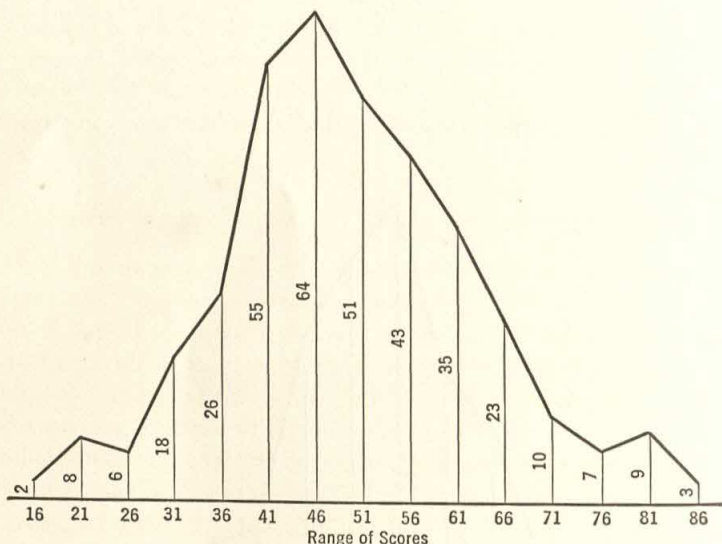


FIG. 48

DISTRIBUTION OF SCORES ON A VOCABULARY TEST MADE BY 360 ELEVEN-YEAR-OLD BOYS

(After Gansl, 1939, p. 27.)

through the tops of these vertical lines. From the figure it appears that only one man in the group had a brain weight of 1,125 gms while 86 had brain weights centering around 1,375 gms. The number of individuals *increases* as we go in from the light to the medium brains; and *decreases* as we go from the medium to the heavy brains. This same arrangement of values may be seen in Figure 48. The "typical" boy with a score of 46 is much more numerous than either the exceptionally poor (score 16) or the exceptionally able boy (score 86).

The distributions of a very large number of human traits follow the general form shown in Figures 47 and 48. Both of these graphs approximate closely the "normal probability curve" shown in Chapter 11, page 227, to illustrate the theoretical distribution of intelligence in the general population. To be sure not all traits in all groups follow the ideal normal curve. The group may be small, heterogeneous, or specially selected (e.g.,

¹ These values actually represent the midpoints of intervals along the baseline.

feeble-minded adults, or children of different ages); the tests may be poorly suited to the group (too easy or too hard); a trait or character may not occur in such a way as to provide a symmetrical distribution. But when groups are large, representative and carefully measured, so many traits have been found to approximate the normal distribution that this ideal curve is used generally in describing individual differences. When a trait is normally distributed, clearly we cannot speak of "types." Every one possesses *some*, at least, of the characteristic under study; and the gradation as we go from small to large amounts of the trait is regular and without breaks.

3. CORRELATION IN THE STUDY OF INDIVIDUAL DIFFERENCES

Suppose we have given two tests—one of memory span and the other of hand-eye coördination—to a group of twelve-year-old boys. Will the boy who ranks highest in memory span rank high, medium, or low in hand-eye coördination? If boys hold the *same* order in the two tests the "correlation" between the two abilities is said to be "high"; if standing in the first test has no relation to standing in the second, the correlation is negligible or zero. The method of correlation is a valuable technique in the study of individual differences. From relationships among measures we can discover which traits and abilities vary together or show correspondence; to what degree individual differences in a variety of measures tend to be consistently maintained; and to what extent related persons, fathers, sons, brothers, sisters, possess traits in the same amount.

The fundamental notions underlying correlation were developed by Galton who used the method in his studies of the inheritance of traits. To Galton belongs the credit of having first shown that quantitative relationships between abilities can be calculated, and of having stimulated interest in such relationships.

Degree of correspondence is measured by the coefficient of correlation—a ratio denoted by the letter r . This coefficient varies from 1.00 (perfect relationship) through .00 (no relationship) down to -1.00 (perfect inverse relationship). Its meaning may be most simply comprehended, perhaps, with the help of such a diagram as that shown in Figure 49. These data, taken from Galton's *Natural Inheritance* (1889), page 208, represent graphically the relationship between the heights of "mid-parents" and their adult offspring. Galton obtained the mid-parent heights by multiplying the mother's height by 1.08 and averaging this figure with the height of the father. The same correction was applied to the heights of the female offspring in order to give the lesser female heights equal weight with those of the male. Each mid-parent is counted separately

with each offspring (of which there were often several); hence the table shows the total number of offspring, namely, 928. The average height for both mid-parents and offspring is close to $68\frac{1}{4}$ inches, and the horizontal and vertical lines drawn through the diagram have been located at these points in the two distributions of height. To illustrate how a "two-way diagram" or correlation table is made up note the entry 34 at the center of the diagram. This entry is opposite the 68.5 inch entry in the mid-parent row of heights and below the 68.2 inch entry in the offspring column of heights. This means that these thirty-four adult offspring with an average height of 68.2 inches possessed parents whose average height was 68.5 inches. The other entries are to be interpreted in the same way. Each offspring has been classified in two ways—with respect to his own height and with respect to his parents' height. It is clear from Figure 49 that the entries tend to cluster in the upper right hand and the lower left hand sections of the diagram. This means that, in general, tall parents have tall children, and short parents short children. In other words, the heights of parents and offspring are related; and this tendency of the two groups to vary together is expressed by a correlation coefficient or r of .45. The calculation of r is somewhat involved, and is described in detail in textbooks dealing with statistical method. The mathematical derivation of r was made by Karl Pearson, an eminent British biometrician. Since Galton's time, literally hundreds of correlations have been calculated, and the relationships of most physical and mental traits are now well established. In a general way correlations above .60 are considered "high"; those from .60 to about .30 "marked" or "moderate"; those from .30 to .00 "low" or negligible. Negative correlations, calculated in the same way as positive relations, imply that larger amounts of the one trait accompany smaller amounts of the other—as when degree of skill or accuracy correlates *negatively* with number of errors made.

The data in Figure 49 may be used to demonstrate another important discovery made by Galton. Table XXXIII shows the median (approximately the average) heights of offspring for given mid-parent heights. Note that the offspring of very tall parents—those 72.5 inches, say—tend to be slightly *below* their parents in height; and this discrepancy holds on down to the mean or average height. Below the mean, offspring tend to be *above* their parents in height (see last column of Table XXXIII). This tendency for the heights of offspring to vary *away* from the heights of their parents and back toward the mean height of the *entire group* (parents and offspring) was called by Galton the *law of filial regression*. This "law" is the expression of a natural tendency which protects the race from extremes. Unless some such principle operated—if tall parents had still taller children, and short parents still shorter offspring—we should

soon have a race of giants on the one hand and of dwarfs on the other. The *law of filial regression* holds in a general way for mental traits as well as for physical characters. Its operation is seen in the relatively infrequent appearance of freaks at either extreme of a distribution of physical or mental traits.

TABLE XXXIII
TENDENCY FOR ADULT OFFSPRING TO "REGRESS" TOWARD
THE MEAN ADULT HEIGHT

(From Fig. 49)

<i>Median Height of Offspring for Given Mid-Parent Height</i>	<i>Height of Mid-Parents</i>	<i>Difference</i>
72.2 inches	72.5 inches	-.3
69.9 "	71.5 "	-1.6
69.5 "	70.5 "	-1.0
68.9 "	69.5 "	-.6
68.2 "	68.5 "	-.3
67.6 "	67.5 "	+.1
67.2 "	66.5 "	+.7
66.7 "	65.5 "	+1.2
65.8 "	64.5 "	+1.3

4. NATURE AND NURTURE IN THE STUDY OF INDIVIDUAL DIFFERENCES

The relative importance of nature and nurture in determining individual differences is an old and much discussed question. Opinion has shifted back and forth from belief in the all-importance of heredity to confidence in the greater potency of environment. A good part—perhaps a major part—of the disagreement here has arisen from a failure to define terms properly. Heredity and environment are not to be thought of as "forces" which operate separately and independently, although many differences among individuals seem attributable primarily to one or the other influence. Thus, the color of a child's eyes and hair, his height, weight, and physique are usually considered to be native or inborn characteristics; while his "mother tongue," e.g., English or French, since it depends upon the country in which he was reared, is held to be a product of environment. But these distinctions are not as clear-cut as they appear to be at first glance. Physical attributes such as hair- and eye-color, height and weight, are affected by climate, diet, and disease; and, without the elaborate structure of the throat and vocal cords, speech of any sort is impossible. In neither case is nature or nurture exclusively responsible.

Every young animal or human child is born with certain skeletal and muscular structures, as well as with a repertory of movements (many highly coördinated) which are native and inherited. But this equipment would be of little value unless shaped by learning and experience. Intrinsic factors such as immediate ancestry, sex, and race always operate in some environmental setting. In any given act, whether it be spelling a word, throwing a baseball, or building a bridge, both native constitution and specific training enter. If behavior were arranged on a scale of complexity, therefore, from the simple reflex to the most complex literary and scientific achievements, there would be no point where heredity left off completely and environment took over. In short, both groups of influences are indispensable to growth and both must be studied by those who wish to understand individual differences.

5. STATISTICAL STUDIES OF INHERITANCE

Galton was much interested in the problem of nature and nurture. In the study of imagery, referred to on p. 266, he concluded that the reported absence of imagery is largely a result of training and of manner of thinking and working. But Galton found considerable evidence, also, that amount of imagery is, at least in part, hereditary. He reports, for example, that imagery tends to run in families; is stronger in women than in men; is stronger in younger than in older persons. Imagery, according to Galton, is not necessarily correlated with eminence as a painter or imaginative writer. Great individual differences exist in the *kind* and *amount* of imagery.

Galton early attacked the question of the inheritance of conspicuous ability or talent. In his first statistical study of the influence of immediate ancestry upon achievement, Galton selected for study 977 eminent British men, each of whom he judged would rank as one man in 4,000 in ability. These selections were based upon a survey of the man's accomplishments, upon his biography, and upon other available information. Among those included in the 977 were judges, statesmen, prime ministers of England from around 1768 to 1868, military commanders, literary and scientific men, poets, artists, and clergymen. Galton's method was to determine whether these men had more eminent relatives—fathers, brothers, sons, uncles, and others—than would be expected of 977 men selected at random from the general population. His findings showed that his selected group of 977 had a total of 332 close relatives just as eminent as themselves as against a probable, or expected, number of only 1. These eminent relatives included 89 fathers, 114 brothers, 129 sons. Moreover, this same group of 977 had 203 equally eminent grandfathers, grandsons, uncles, and nephews, as against a probable number of 3. The probabilities of 1

and 3 here stated mean that the statistical chances are 1 and 3, respectively, that any unselected group of 977 men will have eminent relatives up to Galton's standard of 1 in 4,000. In another study of the inheritance of artistic ability, Galton found that in 30 families wherein both parents were artistic, 64 per cent of the children were artistic, while in 150 families wherein neither parent was artistic, only 21 per cent of the offspring showed artistic ability. Other studies of related individuals revealed a strong tendency for traits to be found in a like degree in parents and offspring. Galton took these data to imply the great superiority of heredity over environment. Even the best environment, he contended, is unable to raise a man to a position of eminence unless he possesses natural gifts of a high order.

The statistical studies begun by Galton were extended and increased in scope by Karl Pearson, Galton's disciple and, from 1911 to 1936, Galton Professor of Eugenics at the University of London. In one typical study by Pearson, two thousand brothers and sisters were rated by their teachers for intelligence, vivacity, conscientiousness, popularity, temper, self-consciousness, assertiveness and handwriting. In each of these characteristics, except handwriting and temper, the individual rated was described simply as falling into the upper or lower group with respect to the given trait. Correlations for all eight traits averaged .52 between brother and brother, .51 between sister and sister, and .52 between brother and sister. Pearson had found previously that the correlations between such attributes as eye-color, hair-color, height, and length of forearm, over which environment can have little if any effect, also average from .50 to .55 for siblings (i.e., brothers and sisters). From these results he argued that since the relationship of physical traits must be due to heredity, and since these characteristics are no more highly related than mental traits, the correlation of the latter must also be basically the result of hereditary influences. He writes (ref. 37, p. 156): "We are forced, I think literally forced, to the conclusion that the physical and psychical characters in man are inherited within broad lines in the same manner and with the same intensity."

These statistical studies carry much weight owing to the eminence of their authors and the care with which they were compiled. But several difficulties arise, nevertheless, to prevent their full acceptance as evidence for the predominant influence of heredity. The first and most obvious difficulty is the error of selection. Galton's choice of eminent men was almost inevitably affected by subjective factors such as personal preference as well as by the relatively greater accessibility of data on some men than on others. It is doubtful whether eminence as a statesman, in which political and social conditions play so large a rôle, can be equated to eminence

as an artist or as a man of science. Again, it is doubtful whether a man's true worth can be correctly judged from the length of his biography or from what his contemporaries think of him. And it is unnecessary to add that we do not know just what part social factors and family tradition, wealth, education, and opportunity played in determining the eminence of Galton's group. The error of selection is almost certainly present, too, in Pearson's ratings. For one thing, the teachers who did the rating knew that the investigation was designed to discover whether children of the same parents resemble each other, and hence very likely were looking for resemblances more often than not. Moreover, pairs of siblings were rated against each other rather than separately on *all* traits, thus giving ample opportunity for likes and dislikes to come into play. This so-called "halo" effect is exhibited when an individual is rated for intelligence, say, and then for reliability, or for honesty, or courtesy; it is well nigh impossible to prevent the first rating from unduly influencing the later ones. Beside the unknown amount of halo in Pearson's ratings, it seems evident that such characteristics as temper, popularity, vivacity, and handwriting must depend to a considerable degree upon health, home conditions, training, and stimulation. It is impossible to say, then, how much of Pearson's fraternal correlation of .50 in mental traits is the result of native and how much of environmental factors. Finally, Pearson's classification of traits into only two categories was extremely rough, so that large errors of placement were highly probable and quite serious.

6. FURTHER STUDIES OF NATURE AND NURTURE

While this early work had many limitations, as we have seen, later studies suggest that Pearson was right in emphasizing the greater potency of heredity over environment. One of the most extensive experimental investigations is Terman's study of gifted children, mentioned elsewhere (p. 229). It may be recalled that a gifted child in Terman's sense meant a child with an IQ of 130 or above. In Terman's experimental group of 644 children, seventy-three families contributed at least two gifted children, and nine families contributed three or more. The number of families in which two bright children were found was more than 1,200 times the number which chance alone would provide. Nearly one fourth of the members of the Hall of Fame were related to these children, while many of their parents and near relatives hold or have held major political offices, college presidencies, professorships, and important business positions. Follow-up studies have shown that these bright youngsters maintained their superior status in comparison with average children.

So much for the power of heredity. On the side of environment, we

find that these gifted children came, in general, from superior homes. Income level was considerably above that of the average American home; the educational level of the parents was high; and the divorce rate low. It can be argued that intelligent parents have intelligent children and make good homes; or that good environment leads inevitably to superior achievement. On biological grounds the argument for heredity appears to be stronger. A not unfair analogy is that of an automobile engine. Good oil and careful handling will undoubtedly make a cheaply built engine run better, but even the best oil and the most careful handling cannot make it perform for long like a first-class machine.

Several experimental studies have attempted to disentangle the contributions of nature and nurture to individual differences in achievement. Jones has reported a study of inheritance in 105 families containing 317 children, born and reared in rural sections of New England. The general intelligence of parents and offspring was measured by the Army Alpha, that of the younger children by the Stanford Revision of the Binet tests. The sample selected for study was particularly adequate, since environmental conditions, education, home training, amusements, and church influences were closely similar for all members of the group. Moreover, all of the individuals tested were native-born whites of old American (mostly British) stock and only English was spoken in their homes. Correlations between the intelligence test scores of brothers and sisters were found by Jones to be .49; of father and child, .51; and of mother and child, .55. These results are very similar—almost identical in fact—with those of Pearson. Working with a group of more than 1,000 siblings, all high school students, Thorndike reports a correlation of .60 between sibling scores on a battery of tests designed to measure general ability. Thorndike remarks that if Pearson's coefficient of .52 for sibling resemblance in physical traits is accepted as giving the undiluted native or inherited resemblance, then the effect of the environment is to raise the sibling correlation for mental ability from .52 to .60. Provided that .52 really represents genetic resemblance, this result assigns an exceedingly meager influence to environment.

Two further studies of the effects of immediate ancestry and nurture may be mentioned before leaving this topic. The first deals chiefly with the influence of environment upon the intelligence of children. It was conducted by F. N. Freeman, K. J. Holzinger, and others at the University of Chicago. The second deals with the relative effects of nature and nurture upon intelligence. It was carried out by Barbara S. Burks at Stanford University under the general supervision of L. M. Terman. Both studies made use of foster children, and accepted the Stanford Revision as their standard measure of intelligence. Foster children were selected for study

because their heredity is usually poor and the foster homes in which they are placed better than their own homes. This gives the investigator an opportunity to discover whether an improved environment will operate to increase IQ. In the Chicago study children who were tested immediately upon placement in foster homes had gained on the average *seven points* in IQ when retested after about four years of residence. Children placed in superior homes (the status of the home was determined by rating devices) had gained about ten points in IQ; those in average homes, about five points. The Stanford study found an average gain of five to six points in IQ as a result of residence in a good foster home. Both studies agree, therefore, that an average home environment may raise the IQ five to six points, while a superior home may increase it still more. Unfortunately this result is not entirely sure, owing to the many uncontrolled factors which enter into the choice of a foster child, uncertainty as to the parentage of many adopted children, and difficulty in defining the cultural status of the foster home. Furthermore, the amount of allowable error in an IQ determination (the "probable error") is about five points. The weight of evidence, therefore, inclines to the conclusion that environment has some but surprisingly little effect upon intelligence.

7. THE STUDY OF TWINS

The most direct way of evaluating the influence of immediate ancestry is through the study of twins, since twins represent the nearest approach to identical heredity. Here again Galton was the pioneer. In an early study he gives an interesting narrative account of his study of twin resemblance. Galton collected data on about eighty pairs of twins from teachers, friends, parents, and the twins themselves. These reports are in the nature of stories, anecdotes, and the like, for the most part stressing the prevalent idea that twins are much more alike in mental characteristics than siblings. Galton's records have little scientific validity, in spite of their historical interest, as reports collected from untrained observers are invariably colored by all sorts of temperamental and emotional inferences. It should be noted, however, that Galton did distinguish between fraternal and identical twins. Fraternal twins are the result of multiple births and develop from two fertilized ova. They may be of the same or of opposite sex. Identical twins probably develop from a single ovum and are always of the same sex. Their resemblance is usually so striking that it is often difficult to tell them apart.

Thorndike's 1905 study was the first quantitative investigation of mental resemblance in twins. He measured fifty pairs of twins on six mental tests and compared the correlation between their scores with the corre-

lation of siblings on the same tests. Correlations for twins ranged from .73 to .85 with a mean at .77, while the coefficients for siblings averaged around .30. Thorndike next divided his twins into *younger* twins (nine to eleven years) and *older* twins (twelve to fourteen years) and computed correlations among his six tests for the two groups separately. The average coefficient for the younger twins was .83, for the older twins .70, indicating that older twins are no more alike than younger twins in the functions measured by the tests. Thorndike argues that if resemblances between twin-pairs in mental ability are the result of common training and common surroundings in school and home, the older twins (because of longer exposure) should be more alike than the younger, especially in those traits much influenced by training. Since his results showed the opposite to be true, he concludes that the evidence favors heredity as the primary cause of twin resemblance.

Many careful studies of twins have been made since Thorndike's pioneer investigation, with more cases, better technique, and better tests; but the results are not very different from his. Merriman found the correlation between IQ (Stanford-Binet) for forty-seven twins, five to nine years old to be .81; for fifty-eight twins, ten to sixteen years old, .76. Lauterbach with twenty-one measures and 212 twin pairs, found like-sex twins (probably identical) to be more alike in both physical and mental characteristics than unlike-sex twins (probably fraternal), the latter being scarcely more alike than siblings. The average correlation for seven mental tests for like-sex twins was .67, for unlike-sex twins .41. Hildreth, applying many measurements, found the average correlation between twins to be approximately .75, between siblings about .50. Leahy, whose study has been mentioned (p. 238), obtained an average correlation of .51 between child and own parents' intelligence and a correlation of .22 for child and foster parents' intelligence. In five motor tasks, McNemar found that identical twins were much more alike than fraternal, the average correlation being .47 for fraternal, and .83 for identical twins. Continued practice in the five motor tasks led to a somewhat closer resemblance between fraternal twins (increase in correlation). But the fraternal twins remained less alike than the identical even after practice.

Fifty identical twins reared together were found to differ negligibly if at all in Stanford-Binet IQ (Newman, Freeman, et al, 34). In the same study, nineteen identical twins separated in infancy or in early childhood and brought up in different families were less alike in IQ than those twins reared together, the greatest discrepancies appearing where the educational and cultural advantages of the twin pairs were most divergent. Feingold has completed a study of 120 twin pairs all between sixty and eighty-nine years of age. The 75 identical twin pairs in this group

were more alike in general intellectual ability than were the 45 fraternal twins. This result is especially significant as showing the persistence of hereditary resemblances in spite of long separation and wide differences in training.

All in all, studies of twins indicate clearly (1) that twins are more alike than siblings; (2) that identical twins are more alike than non-identicals or fraternal; and (3) that environmental influences must operate differentially and markedly before members of a twin pair differ greatly in measured intelligence. The difference between the correlations for mental traits of around .50 (for siblings) and .75 (for twins) may be reasonably attributed to the more nearly identical inheritance of the twin-pairs.

To summarize briefly the last three sections, most careful investigators are convinced that native factors are more potent than environmental in fixing the *initial* amount of an individual's aptitudes and the *extent* to which these may be developed. At the same time, no one believes that potential ability, even when very great, can express itself adequately under a restrictive and circumscribed environment.

8. SEX DIFFERENCES

Competent investigators who have studied the question of sex differences in mental ability are convinced that on the whole such differences are small; and that, when they exist, can more often be attributed to a complex of temperamental and social factors than to innate or hereditary differences in capacity for achievement. The very different training given men and women until recent times; the attitude of conservative people toward women in business or in the professions; the traditional idea of what constitutes a man's work and of what constitutes a woman's work—all of these factors have so colored the question of sex differences in mental ability that an unbiased conclusion is often hard to reach.

One of the earliest studies of sex differences in ability was conducted by Helen Thompson at the University of Chicago in 1903. A variety of tests were employed, including measures of motor ability, sensory acuity and discrimination, and mental ability, the whole battery requiring between fifteen and twenty hours to administer. The subjects were twenty-five men and twenty-five women college undergraduates, 20 to 25 years old. The results of these tests may be summarized as follows: In motor tests of speed and accuracy of movement, except in the sorting of cards according to color, the men did slightly better than the women; in sensory tests of smell, taste, pitch discrimination, and accuracy in judging weights, there were no differences; in tests of rote memory, the women were slightly better than the men; in quickness in solving "ingenuity" problems,

the men were slightly ahead. It must be emphasized that in nearly every case the overlapping was great (see Figure 50) and the differences small, most of them being unreliable. In short, the differences *within* either group were much greater than the differences between typical members

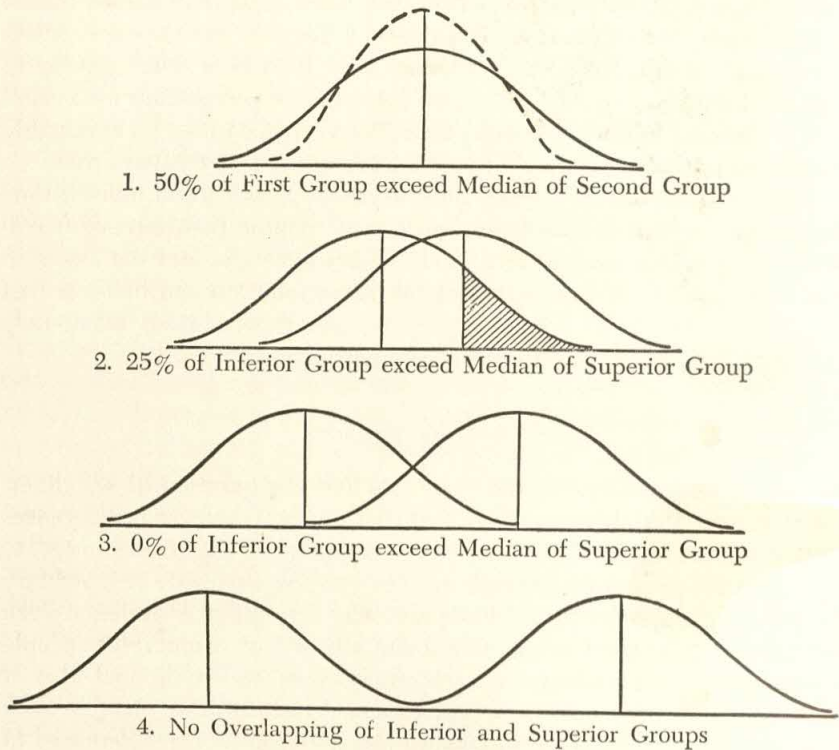


FIG. 50
OVERLAPPING OF FREQUENCY DISTRIBUTIONS

of the two groups: the gap between a high ranking man and a low ranking man was far greater than the gap between the average man and the average woman. Thompson's groups were highly selected, and this fact plus their small size limits the general applicability of her findings.

Pearson, in studying the resemblances of brothers and sisters, had teachers rate these children for such attributes as athletic ability, intelligence, shyness, good temper, and conscientiousness. According to Pearson's findings, boys are more athletic than girls, more noisy, more self-conscious, and quicker tempered; girls are more inclined to be shy, and are more conscientious. Boys and girls were judged by their teachers to

be about equally intelligent. Pearson's ratings, to be sure, are open to those errors of bias and misinterpretation likely to affect all such observations. But in spite of technical faults, his results indicate clearly that in the opinion of their teachers at least, brothers and sisters differ markedly in mental and personal traits.

Hundreds of studies of sex differences have been made for all sorts of traits since these pioneer studies.² It would be impossible, and in many cases not profitable, to report all of these even briefly. We shall try, therefore, to summarize what appear to be the major and best established findings under the following heads: differences in physical characteristics; differences in sensory traits; differences in motor and mechanical abilities; differences in mental and emotional traits. It must always be remembered that when one sex is said to be "better" in some performance than the other, the meaning intended is better "on the average." In all traits there is at least some overlapping of male and female performances (see Fig. 50).

1. *Sex Differences in Physical Traits.* Before puberty, girls are more advanced than boys of the same age in height, weight, dentition, and anatomical and physiological maturity. The brain of the female is smaller than that of the male, but, in proportion to the weight of the body, is heavier. Women have a faster heartbeat than men, and their simple reflexes are usually quicker. Men have larger muscles than women, and surpass women generally in endurance and physical strength.

2. *Sex Differences in Sensory Traits.* Tests in this field have usually shown small differences and often contradictory results. There are probably no real differences between women and men in visual and auditory acuity and discrimination. Women are consistently faster than men in discriminating differences among colors and in quickness of visual perception. Men judge differences in weight somewhat better than women and endure pain better.

3. *Sex Differences in Motor and Mechanical Traits.* Women are consistently faster than men in marking (cancellation) tests. There are no real differences in simple motor functions such as tapping and tracing. Boys are nearly always better than girls in performance tests of a manipulative sort, in completing form-boards, and in tracing mazes. They are also superior in tests of mechanical construction, mechanical aptitude, and knowledge of mechanical things. Much of this difference is probably the result of training and of the early cultivation of different interests. Boys are given mechanical toys and bicycles, girls dolls and fancy wearing apparel, and such things soon come to be identified as belonging peculiarly to one sex.

² See ref. 1 in the bibliography for a comprehensive review.

4. *Sex Differences in Mental and Emotional Traits.* Most studies agree in reporting women as better than men in nearly all tests of memory. Girls almost always do better than boys on vocabulary tests, as well as on tests involving language usage and verbal association. Boys are consistently better in tests involving numbers and spatial concepts (of a geometric sort), and in arithmetical reasoning tests. Men usually do better than women on general information tests, owing partly no doubt to their greater opportunity for contacts in their business and professional lives. In group intelligence tests there are no reliable differences in favor of either sex; the superiority of high school boys over high school girls sometimes reported is probably to be explained by the fact that high school boys form a more highly selected group than high school girls. If dull boys drop out of school earlier and more often than dull girls, or are more often transferred to special classes, those boys who survive to the upper high school grades will, on the average, be somewhat better than their female classmates. Girls usually make better school marks than boys. This has been explained as resulting from the slower physical development of boys, the greater docility of girls, and from various temperamental factors. After selecting his potential geniuses, Terman found that 55 per cent were boys and 45 per cent girls. The occurrence of more very bright boys than very bright girls in the schools has been attributed by some investigators to the greater variability of the male sex in mental traits. Greater variability in males would mean that men range higher (and of course lower, too) on a scale of intellectual ability even when the average man and the average woman are about equally endowed. Greater variability would account for the larger number of male geniuses, and for the superior achievement of men in general. This is a very neat explanation, but it is probably not true, as several careful investigations have found no consistent evidence for greater variability in the male.

Studies of sex differences in emotional and temperamental traits have given few definite and consistent results. Women are reported to be more interested in persons and in personal problems, men in activities and mechanical things. Girls excel boys in tests of moral knowledge and social attitudes, but are no more honest in tests of cheating or deceit. Studies of sex differences in free association indicate a tendency for women's associations to run to personal ornament and to concrete and individual problems; men's associations to business relations, money-making, and more general and abstract matters. Conversation and expressed preferences for books, pictures, music, and the like reflect so many things besides the sex of the person observed that differences found are hardly to be attributed to sex alone. Women have been reported to be more impulsive than men, more introverted, less given to exercising foresight;

but such statements are little more than unsupported pronouncements.

Many studies of sex differences are open to one or more of the following criticisms: (1) Samples are small and are selected by different criteria; (2) tests used to measure abilities are narrow in scope and the meaning of their scores is uncertain; (3) overlap of the groups compared is often ignored, only average differences being reported; (4) the relatively greater maturity of girls age-for-age is not taken into account. Perhaps the most valid conclusion with respect to sex differences in mental ability that we may draw from the available data is that women are *usually* better in tests of language, verbal usage, vocabulary and memory; that men are *usually* better in tests of arithmetic, in performance and manipulative tests, and in the ability to deal with spatial concepts of a geometric sort. Nowhere are the differences large. There are, of course, many considerations which might influence parents and educators in planning different courses of study for boys and for girls. But the assumption of marked sex differences in mental ability should have little weight in determining the choice of school subjects for boys and girls.

9. RACE DIFFERENCES

The study of individual differences arising from remote ancestry or race is complicated even more than is the question of sex differences by prejudice, bias, preconceived notions of superiority and inferiority, and special pleading. There is difficulty, too, in defining just what is meant by a "race" because of the admixture which has gone on all over the world. In general, anthropologists recognize three major racial groupings, the Caucasian or white, the Mongoloid or yellow, and the Negroid or black (a fourth and small group of doubtful classification is often added). Europeans are further subdivided into three groups, Nordic, Alpine, and Mediterranean, on the basis of differences in eye and hair color, stature, head shape, and other anatomical characteristics. Almost every European nation is a compound of all three of these strains. The modern British, for instance, are Scandinavian Nordic, Norman-French Alpine, and Old Celtic Mediterranean, the dominant strain depending upon the section of the country selected for study. The modern Germans and Italians are as mixed a people as the British, the Germans being mostly Nordic and Alpine, the Italian Alpine and Mediterranean. Probably the Swedes represent as pure a strain as any European group, being largely Nordic.

Opinions upon race differences in mentality vary greatly. On the one hand we find a firm belief in the native superiority of certain groups (usually including one's own): on the other, the view that there are really

no hereditary racial differences, such diversities as appear being the result of variations in culture, education and training, traditions and customs, relative isolation from other groups, and climatic and geographical factors. Galton was a firm believer in race differences in native capacity; and it was to improve his own native stock that he initiated the eugenics movement in 1883. Galton believed as a result of his observation and study of the history and accomplishments of different races, reports of travelers, and other data, that the Negro is two grades in mental capacity below the modern Englishman, who in turn is two grades below the ancient Athenian Greek. Galton's scale of intellectual ability embraced sixteen steps ranging from a very high to an extremely low level.

Both Galton's conclusion concerning race differences and his scale for measuring such differences have been criticized, the latter on the ground that it is arbitrary and subjective, the former on the ground that races cannot be fairly compared in terms of number of famous men produced or apparent social and material progress. Admittedly it is hard to define progress, while social and climatic conditions and geographical factors cannot readily be equated. Granted the force of these objections, it would seem nevertheless that wide variations in accomplishment and in the building up of an orderly civilized life must reflect more than differences in environmental pressure. Probably no one would contend seriously that the Hottentots have shown the same capacity for achievement as have the European inhabitants of South Africa, or that Eskimos have shown the same aptitude for scientific invention as have, say, the modern Germans. In other words, Galton's contention that there are native differences in intellect between races far removed on a scale of accomplishment would seem to be justified. As Thorndike aptly remarks (44, p. 221): "Common observation of the African and European, for example, decides that the latter is superior in intellect, enterprise, and self reliance . . . Two races need not be equally gifted because each is equally well adapted to its environment, if the second race has by superior enterprise sought out or created a more exacting but also more remunerative environment . . . The Bushman may count all that he needs to count, but to put one's self in a position that needs algebra and calculus may itself be a symptom of superiority . . . The very fact that a certain test seems unfair to the Bushman may be evidence of his inferiority."

Hundreds of studies have been made in an effort to evaluate the influence of racial origin upon achievement. Without attempting to discuss more than a fraction of these, we shall try in the next few paragraphs to summarize those results which seem most reliably established. Studies of primitive peoples as in the pioneer study of Woodworth made at the St. Louis Exposition in 1904, indicate that primitives do not differ markedly

from the modern American in keenness of vision, in hearing, in sensitivity to pain and pressure, or in delicacy of skin and muscle sense. Form-board tests (p. 137) designed to measure intellectual activity of a simple and rudimentary sort brought out no reliable differences between whites and many less civilized folk, Eskimos, Indians, Filipinos, and Senegalese. But the Negrito and Pygmy (African tribes of small stature and extremely low culture) did no better than low grade or even imbecile whites. Since the feeble-minded white person differs *least* from the normal in physical characteristics and motor abilities and *most* in language and verbal skills (represented by standard tests of intelligence), this result makes it appear probable that the intellectual gap between civilized and primitive folk is a distinct and wide one.

In America many studies have been made of the performance of Negroes upon mental tests, and as result the inferiority of the Negro to the white in mental ability has been frequently asserted. It is hard, however, to tell how much of the Negro adult's apparent inferiority should be attributed to his lower social status, usually inferior training, and lack of opportunity for free social contacts. The fact that the Negro soldier scored lower on the average than the white soldier on the Army tests in both world wars has been discussed elsewhere (p. 257) and some of the difficulties which arise in accepting this finding as evidence of native inferiority have been indicated.

In a study of very young children, McGraw tested sixty-eight white and sixty Negro babies from 2 to 11 months old, with the Vienna "Baby-tests" described on page 136. This study is valuable because the extreme youth of the subjects excluded many social and environmental influences, operative later on, and difficult or impossible to equate. The white babies did better than the Negro on the mental tests at each month level, becoming increasingly better as the upper and more difficult levels of the test were reached. There was, however, considerable overlap in score at each age level, about 25 per cent of the Negro children performing better than the average white baby. The fact that even at this early age the white children tested consistently higher than the Negro strongly suggests heredity as a major cause of the differences in mental ability revealed by the tests. This conclusion is weakened somewhat by the fact that the Negro babies were not as well developed physically as were the white babies and hence were perhaps not as well nourished. In a later study, in which Negro and white babies had much the same diet, differences between the two groups in several measures of physical and sensory-motor development did not appear.

Many studies agree that the average IQ of the Negro child is from 80 to 90; and school surveys have shown that upon educational achieve-

ment tests Negro children test lower grade-for-grade than white children. The gap in score between the two groups becomes progressively larger as we go up the grades, a result which may arise from an increasingly restrictive environment, from an increasing inability of the Negro pupils to deal with the more difficult materials of the tests, or from both causes. In studies of school children, Arlitt and Sunne found fairly consistent differences in mental test performance of Negroes and whites. When Negro and white children were equated for social status, however, the superiority of the white was much reduced. Moreover, within the white group there was a range in IQ of 33 points between children from very superior and those from very inferior social levels. This is a wider intellectual gap than is found between whites and Negroes, but it should be noted that even the "inferior" and "very inferior" whites were slightly ahead of the Negro. This same result was found by Bruce in a rural district in Virginia. White school children achieved higher scores than Negro school children on intelligence tests, even when educational opportunity and social and economic factors were approximately equal. Ferguson administered a number of mental tests to white and Negro school children in three southern cities in an effort to measure abilities less complicated by social factors than are school marks. On these tests the Negroes turned in performances about three fourths as good as those of the whites. When classified roughly as to skin color, Negroes with apparently the highest degree of white ancestry approximated most closely to the white norms.

In an oft-quoted study (Witty), a systematic search was conducted for superior Negro children in seven Chicago public schools in which approximately eight thousand Negro children were enrolled. All told, 103 children were located whose Stanford-Binet IQ's were 120 or above. One fifth of these superior children reported no white forebears, the remaining four fifths being of mixed Negro and white ancestry. The socio-economic status of the homes from which these children came was distinctly above average. This study shows clearly that Negroes are capable of high achievement. At the same time it should be noted that in a sample of eight thousand white children, at least eight hundred (10 per cent) can be expected to possess IQ's of 120 or higher; and this suggests that about eight times as many bright white as bright Negro children will be found in the public schools.

To summarize, on tests of mental ability the American Negro ranks *on the average* consistently lower than the American white. The regularity of this result from babyhood to adulthood makes it appear unlikely that environmental opportunity and social status can explain *all* of the difference found. While only 25 per cent of Negroes do better than the average white, many achieve very high scores. Moreover, the presence of superior

Negro children in the schools is clear evidence that Negro ancestry does not preclude superior achievement.

There have been a few studies of temperamental differences between American Negroes and whites, but none of these is very conclusive. Studies of differences in responses given by Negro and white college students to the items of personality inventories are difficult to interpret because of the many social and economic factors which may operate selectively in the two groups. McFadden and Dashiell found small and unreliable differences between Negroes and whites on tests designed to measure temperamental and impulsive traits. Crane has attempted to compare experimentally the fear reactions exhibited by Negroes and whites in a laboratory situation. His test consisted in measuring the ability to inhibit the withdrawal of the hand from between two uprights when a weight appeared to be falling upon it. Each subject had been told that no harm would come to him, but many were unable, despite this assurance, to keep from pulling the hand away. There were fewer complete withdrawals on the first trial among whites than among Negroes; but a greater *tendency* for the whites to withdraw on subsequent trials, as shown by tentative movements, twitching of the arm muscles and sudden intakes of breath. Crane concluded that Negroes are given to sudden and complete impulsive reactions which once given are over and done with; that whites on the contrary are less likely to react violently at first but are unable to inhibit completely their subsequent reactions.

Studies of American Indians and of Mexicans indicate that these groups generally score lower than American whites on verbal and non-verbal tests of mental ability. There is some evidence that mixed-blood Indians are superior to full-blood, just as mulattos are in general superior in mental test performance to pure-blood Negroes. A part at least of the Indian's inferiority in mental test performance is probably temperamental, the result of a very different attitude toward life as well as of different ideas of what is important and valuable. American Indian children, for instance, while much slower than white children on form-board performance tests, are consistently more accurate; apparently the need for speed has little appeal to the Indian. Orientals, Japanese and Chinese, who have been tested in America are little if at all inferior to native whites, given equal opportunity for education and for contacts with American life.

10. NATIO-RACIAL DIFFERENCES

The comparative showing of various adult groups of European origin upon the Army "combined scale" has been presented elsewhere (p. 260). Table XXXIV gives the results of a comparative study of nearly five thou-

sand school children from 5 to 18 years old. These children, all of whom were offspring of immigrant European parents, were given three group intelligence tests which included both language and non-language items. Results in Table XXXIV are for children grouped according to the nationality of their parents. Comparative scores are expressed in terms of average IQ's derived from the three group tests. Differences are small for adjacent groups, but the range is large as we go from the highest to the lowest. Of the seven highest groups, for instance, two are predominantly Nordic (English and Swedes); two are Alpine (Germans and Lithuanians); two are Alpine and Mediterranean (Polish and Russian Jews); and one is Mediterranean (Irish). Of the seven lowest groups, one is Nordic (British Canadian); three are Alpine (Russian, Polish, and French Canadian); three are Mediterranean (Greeks, Italians, Portuguese). There is little evidence in Table XXXIV of clear-cut natio-racial differences.

Mental tests administered to children of foreign-born parents in New York City and elsewhere agree fairly well that Italian and Polish children

TABLE XXXIV
AVERAGE IQ'S OF CHILDREN OF IMMIGRANT GROUPS

<i>Nationality</i>	<i>No.</i>	<i>Av. IQ</i>
Polish Jews	75	102.8
Swedes	232	102.1
English	213	100.7
Russian Jews	627	99.5
Germans	190	98.5
Lithuanians	468	97.4
Irish	214	95.9
British Canadians	115	93.8
Russians	90	90.0
Poles	227	89.6
Greeks	270	87.6
Italians	350	85.8
French Canadians	243	85.3
Portuguese	671	82.7

born in the United States are consistently below the native-born whites, and that Jewish children are not far below and are often above the performance of the native-born white. In his main group of gifted children (see page 229) Terman found that about 10 per cent were of Jewish extraction—nearly twice the number to be expected from the proportion of Jews in the population covered by his survey. The largest per cent of Terman's group was of British and Scotch ancestry, the percentage of South European stock being very low. This agrees in the main with army test results. We cannot be sure, however, that the parents of Terman's

bright children represented samplings from the same intellectual strata of the foreign countries from which their forebears came. There is some evidence that the immigrants from southern Europe are an inferior sample of their own national population (ref. 9).

On the whole, differences among national groups would seem to be more reasonably explained by selection, differences in training, language, educational background, and culture, than by differences in native ability. This conclusion is supported by the results of a study in which six performance tests were administered to boys 10 to 12 years old in France, Germany and Italy. The Nordic boys were slightly ahead of the Alpine and Mediterranean but the differences were small and non-significant. Test differences among nationalities of the *same* racial stock (e.g., German Nordic and French Nordic) and between city dwelling and country dwelling children were greater than differences among those of unlike racial origin. In short, differences in intelligence measurement *between* national groups appear to be small as compared to differences within either group.

11. SUMMARY ON INDIVIDUAL DIFFERENCES

Perhaps the most salutary impression which the reader can carry away from a survey of studies of individual differences is a better appreciation of the difficulties encountered in identifying and evaluating the factors which lead to variations in performance among persons and groups. While the study of individual differences has progressed far beyond Galton's early work, it was Galton's invention and use of statistical and experimental techniques and his insight in seeing problems that gave the impetus to later investigation. To quote E. G. Boring, Galton was "the father (in large part) of mental measurement of individual differences with respect to traits . . . and the originator of the questionnaire and the theory of eugenics." As a pioneer in these fields, Galton's place in the history of science is assured.

BIBLIOGRAPHY

1. Anastasi, A., & Foley, J. P. Differential psychology (revised). New York: The Macmillan Company, 1949.
2. Arlitt, A. H. On the need for caution in establishing race norms. *J. appl. Psychol.*, 1921, 5, 179-183.
3. Betts, G. H. The distribution and functions of mental imagery. *Teachers Coll. Contrib. Educ.*, 1909, No. 26.
4. Bruce, M. Factors affecting intelligence test performance of whites and Negroes. *Arch. Psychol.*, New York, 1940, 36, No. 252.
5. Burks, B. S. The relative influence of nature and nurture upon mental

- development. Twenty-seventh Yearbook of the National Society for the Study of Education, 1928, 219-316.
6. Cattell, J. McK., & Farrand, L. Physical and mental measurements of the students of Columbia University. *Psychol. Rev.*, 1896, 3, 618-648.
 7. Crane, A. L. Race differences in inhibition. *Arch. Psychol.*, New York, 1923, 9, No. 23.
 8. Feingold, L. A psychometric study of senescent twins. Ph.D. thesis. Columbia University. (To be published in 1952.)
 9. Franzblau, R. N. Race differences in mental and physical traits studied in different environments. *Arch. Psychol.*, New York, 1935, 26, No. 177.
 10. Ferguson, G. O. The psychology of the negro. *Arch. Psychol.*, New York, 1916, 5, No. 36.
 11. Freeman, F. N., Holzinger, K. J., & Mitchell, B. C. The influence of environment on the intelligence, school achievement, and conduct of foster children. Twenty-seventh Yearbook of the National Society for the Study of Education, 1928, 103-216.
 12. Galton, F. Hereditary genius. London: Macmillan, 1925. First published in 1869.
 13. Galton, F. Inquiries into the human faculty and its development. London: Macmillan, 1883.
 14. Galton, F. Natural inheritance. London: Macmillan & Co., Ltd., 1889.
 15. Gansl, I. Vocabulary; its measurement and growth. *Arch. Psychol.*, New York, 1939, 33, No. 236.
 16. Gilbert, J. A. Researches on the mental and physical development of school children. *Yale Psychol. Lab. Studies*, 1893, 40-100.
 17. Griffiths, C. H. Fundamentals of vocational psychology. New York: The Macmillan Company, 1924, chapters 2, 3, 4, 13.
 18. Hildreth, G. H. The resemblance of siblings in intelligence and achievement. *Teachers Coll. Contrib. Educ.*, 1925, No. 186.
 19. Hirsch, N. D. M. A study of natio-racial mental differences. *Genet. Psychol. Monog.*, 1926, 1, 231-406.
 20. Hunter, W. S., & Sommermier, E. The relation of degree of Indian blood to score on the Otis intelligence test. *J. comp. Psychol.*, 1922, 2, 257-277.
 21. Intelligence: its nature and nurture. Thirty-ninth Yearbook of the National Society for the Study of Education, 1940, Parts I & II.
 22. Jones, H. E. A first study of parent-child resemblance in intelligence. Twenty-seventh Yearbook of the National Society for the Study of Education, 1928, 61-72.
 23. Klineberg, O. An experimental study of speed and other factors in racial differences. *Arch. Psychol.*, New York, 1928, 15, No. 93.
 24. Klineberg, O. A study of psychological differences between "racial" and national groups in Europe. *Arch. Psychol.*, New York, 1931, 20, No. 132.
 25. Kraepelin, A. Clinical psychiatry. Abstracted and adapted by A. R. Diefendorf. New York: The Macmillan Company, 1918.
 26. Lauterbach, C. E. Studies in twin resemblances. *Genetics*, 1925, 10, 528-568.
 27. Leahy, A. M. Nature-nurture and intelligence. *Genet. Psychol. Monog.*, 1935, 17, 235-308.
 28. Lincoln, E. A. Sex differences in the growth of American school children. Baltimore: Warwick & York, 1927.

29. Mayo, M. J. The mental capacity of the American Negro. *Arch. Psychol.*, New York, 1913, 4, No. 28.
30. McFadden, J. H., & Dashiell, J. F. Racial differences as measured by the Downey Will-Temperament test. *J. appl. Psychol.*, 1923, 7, 30-53.
31. McGraw, M. B. A comparative study of a group of southern white and negro infants. *Genet. Psychol. Monog.*, 1931, 10, 1-105.
32. McNemar, Q. Twin resemblances in motor skills, and the effect of practice thereon. *J. genet. Psychol.*, 1933, 42, 70-99.
33. Merriman, C. The intellectual resemblance of twins. *Psychol. Monog.*, 1924, 33, No. 5.
34. Newman, H. H., Freeman, F. N., & Holzinger, K. J. Twins: a study of heredity and environment. Chicago: University of Chicago Press, 1937.
35. Pasamanick, B. A comparative study of the behavioral development of Negro infants. *J. genet. Psychol.*, 1946, 69, 3-44.
36. Pearl, R. Medical biometry and statistics. New York: W. B. Saunders Company, 1930.
37. Pearson, K. On the laws of inheritance in man. II. On the inheritance of the mental and moral characters in man, and its comparison with the inheritance of the physical characters. *Biometrika*, 1904, 3, 131-190.
38. Rugg, H. Is the rating of human character practicable? *J. educ. Psychol.*, 1921, 12, 425-438, 485-501; 1922, 13, 30-42, 81-93.
39. Stern, W. *Die differentielle Psychologie*. Leipzig: J. A. Barth, 1921.
40. Sunne, D. Comparison of white and negro children in verbal and non-verbal tests. *Sch. & Soc.*, 1924, 19, 469-472.
41. Terman, L. M. *Genetic studies of genius*. Stanford University Press, 1925, Vol. I.
42. Thompson, H. B. *The mental traits of sex*. Chicago: University of Chicago Press, 1903.
43. Thorndike, E. L. Measurements of twins. *Arch. Phil. Psychol. & Sci. Metho.*, 1905, No. 1.
44. Thorndike, E. L. *Educational psychology*. New York: Teachers College, Columbia University, 1925, Vol. III, 220-221.
45. Thorndike, E. L. The resemblance of siblings in intelligence. *Twenty-seventh Yearbook of the National Society for the Study of Education*, 1928, 40-53.
46. Thorndike, E. L., Lay, W., & Dean, P. R. The relation of accuracy in sensory discrimination to general intelligence. *Amer. J. Psychol.*, 1909, 20, 364-369.
47. Witty, P. A. Intelligence as related to race. 39th Yearbook, NSSE, 1940, chapter 12, 257-269.
48. Witty, P. A., & Jenkins, M. D. The educational achievement of gifted Negro children. *J. educ. Psychol.*, 1934, 25, 585-597.
49. Woodworth, R. S. Race differences in mental traits. *Science*, 1910, 31, 171-186.
50. Woodworth, R. S. *Experimental psychology*. New York: Henry Holt and Company, 1938, 39-47.

VISUAL AND AUDITORY PERCEPTION: THE WORK OF HELMHOLTZ AND OTHERS

IN THE history of experimental psychology, no topic has excited more research, and led to more discussion than that of how we perceive or come to know the positions, relations, and meanings of things around us. Reasons for this interest are readily apparent. Sense impressions lie at the basis of all mental activity, and have an enormous practical importance in everyday life. The two senses best fitted for the perception of objects in space are touch and vision, since both the skin and the retina are spread out spatially in two dimensions. Of these two, vision is easily the most valuable for everyday life. Second only to vision is audition. Seeing and hearing, in fact, are usually called the "higher" senses in contrast to tasting and smelling, the "lower" senses. This distinction is due not so much to the greater complexity of eye and ear, as to the relatively greater value of the sense data derived from them. To be sure, a man who loses his taste or smell is decidedly handicapped; but such deprivation is not comparable to the loss of sight or of hearing.

In the following sections we shall first describe briefly the physical apparatus concerned in hearing and seeing. Important experimental findings in the field of auditory and visual perception will then be presented.

1. THE AUDITORY APPARATUS

While it is true that in a general way "we hear with our ears," the most indispensable part of the auditory apparatus is the inner ear. The outer ear acts as a collector of sounds, and as a not too efficient ear-trumpet. The middle ear transmits the sound waves, which impinge upon the ear drum, to the inner ear. From the inner ear nerve impulses pass over the auditory nerve to the brain—and we hear sounds. An important part of the inner ear is a tissue-like structure called the basilar membrane—upon which rest extremely sensitive receiving cells. The basilar membrane, which is often compared to the strings of a piano, is made up of some 24,000 fibers ranging in size from short and thick to long and thin. These

fibers and their sensitive cells constitute the essential auditory receptive apparatus.

If one throws a stone into water, concentric rings will surround the splash. These "rings" grow larger and larger as they move out in wave-like form. In the same manner, sound waves arise as vibrations in the air or in any ponderable medium. Most of the sounds we hear result from the back-and-forth movement of air particles. But sound is also transmitted in other ways, as we all know; through metals, through the ground, and through water. The Indian put his ear to the ground to hear distant sounds, and the swimmer knows how loud the clapping of two stones can sound under water.

Sounds differ in pitch, in loudness or intensity, and in timbre or tone quality. Variations and differences in these characteristics enable us to tell quickly, and usually accurately, one voice from another over the telephone—when the voice is the only cue to the person's identity. The pitch of a musical note depends directly upon the vibration rate of the sound producing body—string or resonance chamber (wood-wind instrument). The high squeaky voice and the deep bass rumble differ markedly in pitch as well as in intensity and quality. Loudness or intensity varies with the amplitude or size of the sound waves, but judgments of intensity also depend upon pitch or vibration rate. As a tone dies away, it becomes lower although its vibration rate does not change. High tones, therefore, are said to be "intrinsically" louder than low tones. Pitch differences from 25 vibrations per second or less to 20,000 vibrations per second or more can be heard by the human ear.

Timbre or tone-quality is an important differentiating characteristic; it depends upon the wave composition of the sound. If two tones which have much the same pitch, 256 and 260 vibrations per second, say, are sounded together, *beats* arise—their number depending upon the differences in vibration rate between the two tones. Beats are often heard as disagreeable variations in intensity, and sometimes as rattles or burrs accompanying the tones. When the difference in vibration rate between the two beating tones is 40 or more, the beats disappear and a new low tone or "difference" tone appears as part of the tone complex. Besides difference tones, other tones with vibration rates which are multiples of the fundamental also accompany a single tone. The difference between the shrill sharp note of the piccolo, and the rich full note of the violin resides in the number and character of the accompanying "overtones."

The best instruments constructed by psychologists for gauging one's "ear for music" or ability to perceive differences in tone are the Seashore tests of musical talent. These tests consist of six double disk records which can be played on any standard phonograph. Each record is designed to

cover a different aspect of musical aptitude: pitch discrimination, intensity, time, rhythm, consonance, and tonal memory are the characteristics measured. Three difficulty levels of these tests are provided. Series A is for people in general; Series B for musically educated groups; and Series C is a precise test for individual use. Seashore's tests have been carefully standardized, their administration is not difficult, and the scoring is objective. In the test for pitch discrimination, for example, a person must decide, when two tones are presented in succession, whether the second is higher or lower than the first. There are 100 comparisons in all ranging from very easy to very difficult. Similarly, the tests of time, intensity, and rhythm demand judgments by the observer based upon a series of comparisons. The test of consonance is somewhat different from the other tests. The observer is no longer called upon to give a judgment which can be checked against known physical differences in the stimuli. Instead, he must decide which of two tone-combinations is the more pleasant, i.e., shows greater fusion, less harshness and noise effects. The test of tonal memory requires the observer to tell which of two, three, or four tones played in succession has been changed upon a second presentation. Tests like these when used in conjunction with measures of ability to sing and read music are useful in the selection of persons who possess potential musical talent.

2. HOW SOUNDS ARE LOCATED IN THE ENVIRONMENT

Not only do we identify sounds as the ringing of an alarm clock or the conversation between two friends, but we locate them somewhere in space. Sound localization, especially in familiar surroundings, is usually immediate and fairly accurate. This is shown by the fact that we are disturbed when we cannot locate a given sound at once. It is hard to see, however, how such a receiving apparatus as the ear can supply exact knowledge of distance and direction. And for this reason, psychologists have spent much time on the problem of how we localize sounds in our environment.

Experiments on sound localization are usually carried out by means of a "sound cage"—a metal circle supported by uprights. Such an apparatus is shown in Figure 51. The subject is seated blindfolded in the center of the circle and sound stimuli—usually neutral sounds like the click or snap of a telegraph snapper—are presented from different points in the cage. The subject locates the position of the sound by pointing or when he has learned the angular divisions on the circle by giving positions in degrees.

Certain definite results have been obtained by use of the sound cage.



Bettmann Archive

HERMANN LUDWIG F. VON HELMHOLTZ

1821-1894

Sounds to right or left are most sharply distinguished—the difference in relative intensity of the sound as heard by the right and left ear being apparently the primary factor here. Sounds directly ahead or directly behind the observer are localized fairly accurately; better localization is obtained when the observer is allowed to turn his head. Sounds above or below the observer are in general poorly placed. We all know how difficult it is to locate an airplane—the sound cannot be “anchored” to anything familiar (see Fig. 52). In ordinary hearing, cues furnished by vision and by past knowledge of the sound (a familiar voice or the opening of a door) are helpful and perhaps crucial in effecting localization. Blind

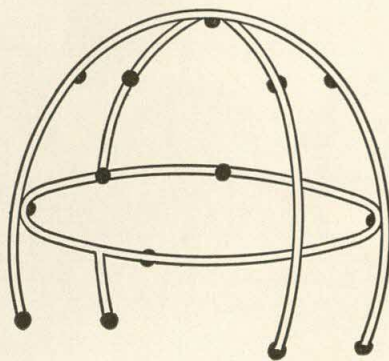


FIG. 51
SOUND CAGE

The horizontal axis is placed at the level of the subject's ears. Stimulus points (black circles) are possible in horizontal, median, and transverse planes and above, below, or at the level of the ears.

persons depend more upon echo, resonance, and secondary cues than do the seeing. This is shown by the fact that persons deaf or partially deaf in one ear locate sounds more accurately when in the sound cage than do normal people with one ear plugged. One-ear hearing is, however, definitely less efficient than two-ear hearing.

An interesting experiment has been made in which sounds entering the two ears were interchanged. Tubes fitted into the two ears were carried over the head to a trumpet or receiving horn just over the opposite ear. In this way each ear got the stimulus which would ordinarily have affected the other ear. The observer wore this instrument (called a pseudophone) an hour or so a day for two weeks and then all day for three days. At first there was general confusion. A horse would be seen trotting along a street and the sound of his hoofs would come from the opposite direction. A person coming through a door would be heard from the opposite side of the room. When the eyes were closed the reversed localization

persisted without improvement throughout the experiment. When the eyes were open, however, correct placement of sounds was never completely lost and got steadily better as the experiment proceeded. More and more the visual cues dominated the auditory, so that by the end of the experiment the observer was localizing fairly well. When the pseudophone was removed the return to normal localization was immediate.



FIG. 52

DIRECTION FINDER

The operator sits in the apparatus and aims himself at the sound. The pair of horizontal artificial "ears" are for horizontal directions; the vertical pair, for direction in height. (From *Psychology for the Armed Services*, 1945, Infantry Journal.)

3. THEORIES OF HOW WE HEAR

Several theories have been proposed to explain the various phenomena of hearing. The oldest and best known is the resonance or "piano" theory of the great German scientist, Hermann von Helmholtz (1821-1894). This theory holds that the long and short fibers of the basilar membrane pick up different sound waves like the strings of a piano. The vibrating membrane agitates the sensitive cells resting upon it, and these in turn pass the effects of the stimulus on to the auditory nerve. Differences in length and in thickness of the basilar membrane fibers enable us to distinguish sounds of different vibration rate, and hence to hear high, low, and intermediate tones.

Another theory, called the frequency or "telephone" theory, suggests that the basilar membrane vibrates as a whole rather than by parts. Nerve

impulses thus set up are thought to preserve faithfully the sound waves so that their vibration frequency is exactly reproduced in the brain center. Instead of a complex tone being analyzed into its parts in the ear itself, as the piano theory supposes, the telephone theory holds that analysis takes place in the brain.

Recent experiments have modified the telephone theory. In one careful experiment, Wever and Bray exposed (under anesthesia) the auditory nerve of a cat from the point where it leaves the inner ear to where it enters the brainstem. When a tone of known composition was sounded at the cat's ear, it was discovered that the sound came through the auditory nerve without loss or change in vibration frequency. Thus it might seem that the frequency rate of a sound is accurately conveyed to the brain as the telephone theory supposes. But further objection has been offered to the telephone theory on the ground that the auditory nerve cannot transmit the vibration rates of very high tones. Nerves do not carry more than 1,000 impulses per second as a theoretical maximum and the upper limit of human hearing is around 20,000 vibrations per second. To meet this difficulty Wever and Bray have proposed the "volley theory," an interesting and ingenious modification of the telephone theory.

According to the volley theory the separate nerve fibers respond differentially to the successive sound waves set up when a note is heard. To illustrate, suppose a tone of 10,000 vibrations is received in the ear. Instead of each single nerve fiber responding to each vibration, according to this theory nerve fiber #1 responds, say, to every *tenth* vibration, beginning with the first—e.g., 1, 11, 21, 31, 41, etc. In like manner, nerve fiber #2 responds to every tenth vibration, e.g., 2, 12, 22, 32, 42, etc. Nerve fiber #3 responds to every tenth vibration, 3, 13, 23, 33, etc. In this way the nerve fibers discharge by groups or squads instead of all at once. No fiber carries more than 1,000 impulses per second, but all together carry 10,000 vibrations per second without "over-loading" the nerve.

While no theory accounts fully for all of the facts of hearing, the volley theory is at present the most comprehensive. Much experimental work is still being done.

4. VISUAL PERCEPTION OF DISTANCE AND DEPTH

The immediate stimuli to visual perception are light-rays which reach the two retinas (the inner surfaces of the eyes) by way of an aperture called the pupil. Figure 53 gives a simple schematic representation of the mechanics of this process. Although the retina is a slightly concave surface, a two-dimensional "field," yet, as we all know, our perceptions of distance, depth, and motion are given in three dimensions. How, we may ask, is this possible? Why aren't the objects of our experience--

houses, trees, and people—spread out before us as though pinned upon a large flat board parallel to our retinas? This has been a hotly debated question in psychology, and much of the experimental data on the subject go back to the researches of Helmholtz. Besides numerous studies dealing with the physiology of the eye and the physics of optics, Helmholtz accumulated many data of direct psychological interest upon the question of space perception. It is largely upon such evidence that the *empirical theory* of perception is based. This theory holds that our knowledge of distance and depth is built up largely, if not entirely, through the mutual coöperation and checkup of vision, touch, and locomotion. The *nativistic theory*, on the other hand, explains facts in three dimensions as natively given and unlearned. Evidence can be marshaled to support this theory.

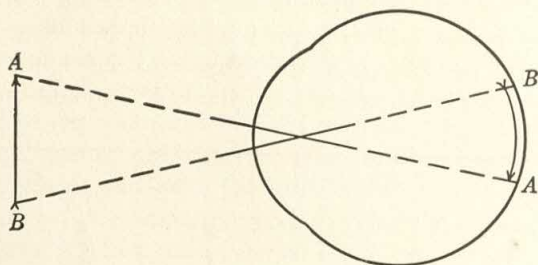


FIG. 53

HOW THE IMAGE OF AN OBJECT AB IS FOCUSED ON THE RETINA

For example, careful experimental work has been done upon the ability of the eye to adapt to light and dark illumination and to color changes as well as to perceive objects in space. It has been found that single retinal nerve cells, when stimulated under different conditions, give curves of response which resemble those of the eye itself. There is no need, then, to go beyond peripheral factors in order to explain the nature of sense organ response. Probably most psychologists today are inclined to accept the empirical view with the strong reservation that there must be a substantial framework to visual perception which is natively given (see Chapter 4).

5. HOW WE SEE OBJECTS IN THREE DIMENSIONS

The invention of the reflecting stereoscope by Wheatstone in 1833 permitted a direct experimental attack upon the problem of how we see objects in three dimensions. Wheatstone's stereoscope was a scientifically accurate instrument, but somewhat clumsy and impracticable for ordinary work. It has been generally displaced by the refracting stereoscope invented by David Brewster somewhat later (*circa* 1843). Essentially, the

stereoscope is an optical instrument by means of which two plane pictures may be presented, the one to the right eye and the other to the left, to give a single fused picture in three dimensions. This striking result follows from the fact that man is a two-eyed animal. His eyes are separated spatially, so that when they are focused upon a common object, each eye views it from a slightly different position. In other words, two pictures are presented in binocular vision, and it is the fusion of these two which gives us our visual perception of three dimensions.

Simple stereoscopic effects may be obtained in the following way: Hold a pen and a pencil straight out in front of the face, about one foot away, and a little to the left of an imaginary line passing through the nose. Now move the pencil slightly to the right of the pen and about two inches away from the face. Close the right eye and draw on a sheet of paper the two objects as they appear to the left (open) eye. Now close the left eye and draw the two objects as they appear to the right eye. It will be sufficient to represent the two objects simply by straight lines. Your pic-

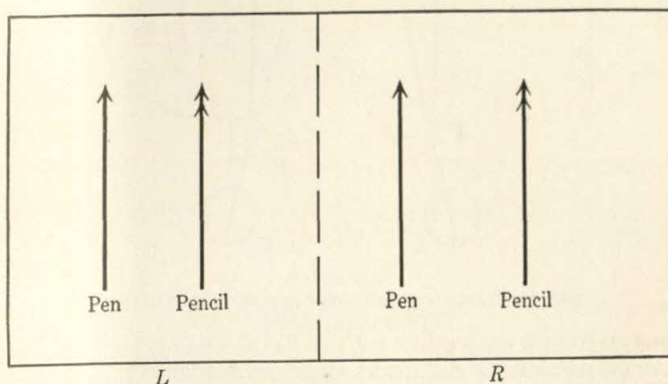


FIG. 54

STEREOSCOPIC SLIDE DRAWN TO GIVE PERCEPTION IN THREE DIMENSIONS

ture should now look like Figure 54, in which the single-tufted arrow represents the pen and the double-tufted arrow the pencil. Be sure and separate the left-eye and right-eye pictures by about two and a-half inches in your drawing. Now hold the completed picture about twelve to eighteen inches in front of the face. Focus the eyes as though you were looking at a far-distant point, and keeping this fixation steadily, move the figure in slowly until the right-hand and left-hand pictures fuse, and there is only one pen and one pencil. When this occurs, the pen will seem nearer than the pencil.

A simple explanation of this interesting phenomenon derives from the

fact that the distance between the left and right pencils is greater than the distance between the left and right pens (this can easily be verified from your picture). Hence, the eyes must converge more from the parallel position to make the pens combine than to make the pencils combine. The greater the convergence, the nearer the object, and so we see the pen as nearer than the pencil. It is a common experience that near objects require greater convergence to be brought to a focus than far objects; and so convergence is a direct cue to relative distances.

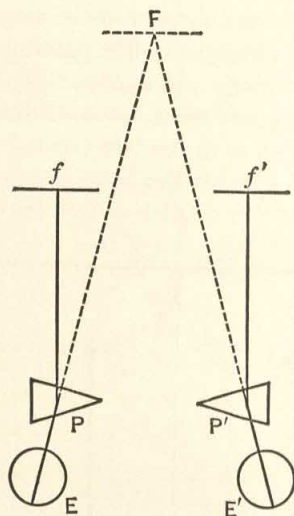


FIG. 55
THE PRINCIPLE OF THE REFRACTING STEREOSCOPE
(Brewster's Model)

The picture for the left eye is placed at f ; that for the right eye at f' . Refraction of the light rays by the prisms at P and P' gives a single fused image at F .

In Figure 55, the principle upon which the refracting stereoscope works is represented schematically. The picture seen by the left eye is placed at f , and the picture seen by the right eye at f' . Owing to refraction of the light-rays by the prisms P and P' , the two pictures are combined by the two eyes E and E' to be seen as a single fused image at F . The impression of distance and depth got from pictures combined stereoscopically is astonishingly real. In taking pictures for ordinary stereoscopic work, the camera first takes the right-eye picture; then it is moved about three inches to the left to take the left-eye picture. Two and one-half to three inches is approximately the interocular distance. Because of this separation, each eye, as we have said, sees around different sides of objects, the right eye around the right, the left eye around the left, and in this way cues to distance and depth are secured. If two pictures are taken, separated

by more than three inches—by one foot, for example—striking and often ludicrous results will be obtained when they are combined stereoscopically. On such a slide we see objects as we would see them if our eyes were a foot apart. Seeing around corners, therefore, becomes highly exaggerated, and objects and people are stretched out in amusing proportions.¹

An illustration of the ordinary Brewster stereoscope is shown in Figure 56. The hood serves to exclude from the eye nearly all of the light which does not come directly through the prisms. The partition between the two prisms restricts each picture to the eye for which it is intended, while

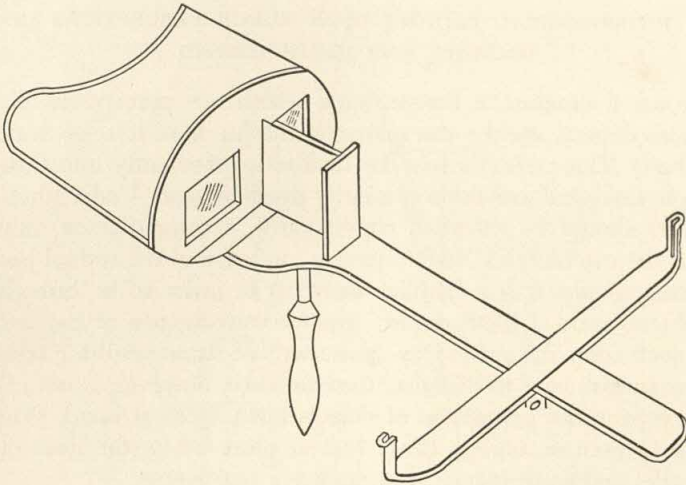


FIG. 56
BREWSTER'S HAND STEREOSCOPE

the extension slide permits adjustment for eyes of different focal lengths.

The part played by the right-eye and left-eye images in determining our perceptions of objects in three dimensions is even more strikingly shown by the pseudoscope than by the stereoscope. In the pseudoscope, the pictures belonging to the right and left eyes are transposed, so that the right eye sees the left eye's picture, and vice versa. The net result of this is a complete reversal of perspective in the fused picture; near things seem far, and far things seem near. Go back for a moment to Figure 54. If this slide is cut vertically into two equal parts and the parts transposed, the pencil pictures will now be closer together than the pen pictures. Hence, when the right and left pictures are combined, more convergence is required to combine the pencils than the pens, and ac-

¹ Titchener's Stereoscopic Slides (Stoelting), forty-three in number, include several in which the interocular distance is much increased.

cordingly the pencil is judged to be nearer. In simple line drawings and geometrical figures, pseudoscopic effects (reversal of perspective) can usually be secured readily enough, either by transposing the right-eye and left-eye pictures, or by inverting the whole slide. With actual scenes, landscapes, objects, and people, however, it is difficult to get pseudoscopic effects when the result would be meaningless, or contrary to experience. Few people, for instance, can see a human face pseudoscopically, i.e., concave, as the "central" or associative facts of experience are too well established. The reversal of lifetime habits is not readily achieved.

6. PHYSIOLOGICAL FACTORS UPON WHICH PERCEPTIONS OF DISTANCE AND DEPTH DEPEND

There are a number of factors upon which our perceptions of three dimensions depend besides this matter of the fusion of left-eye and right-eye pictures. These criteria may be divided conveniently into those primarily physiological and those primarily psychological. Under physiological factors should be included convergence, accommodation, muscular strain on the eye muscles, double images, and size of the retinal image.

1. *Convergence*. It is a familiar fact that in order to be brought into focus objects near at hand require greater convergence of the two eyes than objects some distance away. Sensations of strain resulting from convergence furnish cues to distance. Convergence, however, is not of much value except in the perception of objects fairly close at hand. When the eyes are focused on objects thirty feet or more away, the lines of sight are parallel and convergence is no longer a real factor.

2. *Accommodation* to distance is effected by the eye through changes in the crystalline lens. The lens is a translucent prism-like structure which lies behind the pupil and between it and the retina. Light-rays coming through the pupil are refracted by the lens, and focused upon the retina. When a photographer wants to adjust his camera for distance, he lengthens or shortens the distance between the lens and the screen. The eye, on the other hand, adjusts to distances by increasing or decreasing the thickness (refractive power) of the lens. This pull, or release, of the muscles controlling the lens for different degrees of accommodation, supplies us with cues to distance and depth. But since far objects require little accommodation, this factor, like convergence, is limited mainly to objects close by.

3. *Muscular strain* on the eye muscles. Each eye is moved by six little muscles arranged in pairs. When we compare the lengths of two lines, or gauge the width of a creek or the height of a building, the degree of strain on the eye muscles is a valuable factor in determining our judgment of distance.

4. *Double Images*. The doubling-of-images factor (though mainly physiological) is also of psychological interest because it is so seldom noticed in everyday life. In fact, Helmholtz remarked that the existence of double images "remains unknown to many people even in adult life." Their existence may readily be shown, however, in the following simple way. Focus on a point on the opposite wall some ten to fifteen feet away, and then (holding this fixation) bring a pencil up directly before the eyes and about twelve inches away. The pencil will appear double—unless your focus slips—and Figure 57 shows why this is true. When you focus upon F , the pencil images, since they fall upon non-corresponding

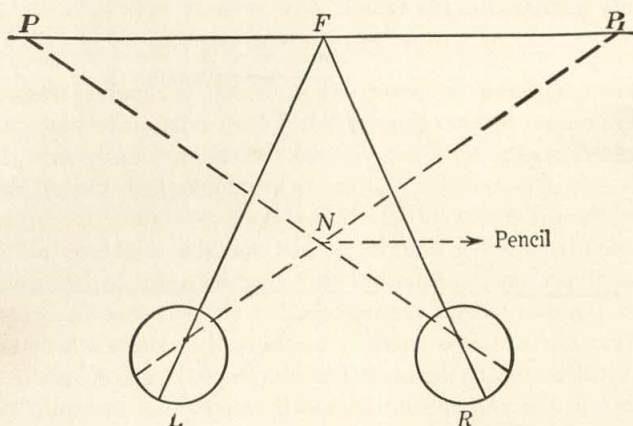


FIG. 57

THE DOUBLING OF "NEAR" OBJECTS IN FAR FIXATION

L , left eye; R , right eye; N , near point; F , far point. When the eyes are focused at F , the pencil at N is seen as two pencils, at P and P_1 respectively, in the plane of F .

retinal points (points which do not give single vision), appear as two pencils at P and P_1 in the plane of fixation. If you close the left eye, the pencil image (seen now by the right eye) is on the left side; and if you close the right eye, the pencil image (seen by the left eye) is on the right side. That is, the images are crossed, and hence crossed images always mean an object *nearer* than the fixation point. If the focus is upon the near object, it is the far object that doubles, as you can readily demonstrate by focusing upon your pencil (twelve inches away) and doubling a picture on the wall in the same line of vision. In this case the double images are not crossed, as Figure 58 shows, and hence uncrossed images always mean an object farther away than the fixation point.²

The essential facts in Figures 57 and 58 may be more easily recalled, perhaps, with the help of the couplet proposed by Titchener:

"Remote regard reverses
Nearer, notice not"

5. *Size of Retinal Image.* From Figure 53 it is obvious that at the same distance from the eye a *large* object will subtend a greater visual angle than a *small* object, and that the retinal images will vary proportionally. Also it is clear that the farther away an object is from the eyes, the smaller its retinal image. Retinal images furnish valuable cues as to distance and size, which are supplemented, of course, by experience and by other sensory data. The man a block away takes up much less retinal space in your eyes than the ink bottle on your desk, but the two objects are interpreted as being at different distances from the eyes, each its natural and customary size. The retinal images vary both for size and distance, and are only partial cues for either.

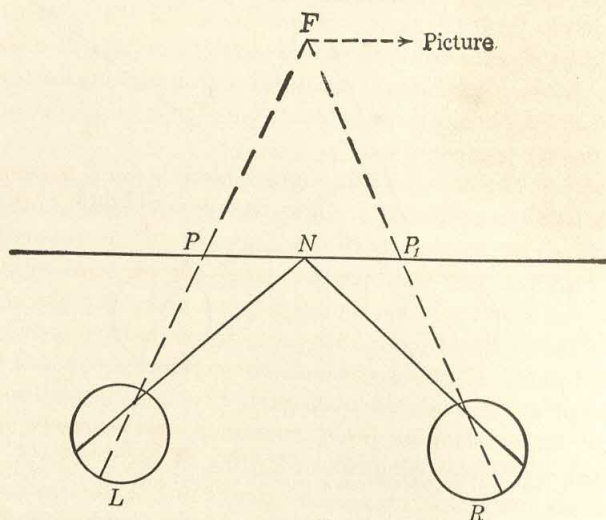


FIG. 58

THE DOUBLING OF "FAR" OBJECTS IN NEAR FIXATION

L, left eye; *R*, right eye; *N*, near point; *F*, far point. When the eyes are focused at *N*, the picture (or other object) at *F* is seen double, at *P* and at *P'*, in the plane of *N*.

7. PSYCHOLOGICAL FACTORS IN DISTANCE AND DEPTH PERCEPTION

Just as important as the physiological factors in space perception are the so-called psychological factors. These depend upon learning and are the result of everyday contacts with near and far objects. Among such determiners may be mentioned superposition, clearness of outline (aerial perspective), shade and shadows, relative movement, and some of the Gestalt effects.

1. *Superposition.* This factor is of value in relative distance judgments.

When a tree, for instance, partially obstructs our view of a house, the immediate inference is that the tree stands *between* us and the house, and hence must be nearer.

2. *Clearness of outline* generally means nearness; dimness and blurring, distance. On a very clear day we are often astonished by the apparent nearness of distant objects, while individuals accustomed to fog and an obscured sky are often completely misled by distances at a high altitude where the moisture content of the atmosphere is very low.

3. *Differences in lighting, shadows, and shade* are especially valuable cues in the perception of depth. Brightly lighted objects in the visual field are generally higher than those dimly lighted, since light usually comes from above; again, on a plane surface depressions or concavities are darker, while protrusions or convexities are brighter. Artists depend to a considerable degree upon light and shade to represent distances and depth. Near objects seem "down" in the foreground, and far objects "up" closer to the horizon, and these facts are regularly employed in pictures to give the proper perspective.

4. *Relative movement* is an important factor in space perception. Objects move toward us and away from us, before and behind other objects, and in this way give us cues as to relative position and relative distance. We are all familiar with the fact that when we are moving, on a train for instance, far objects move along with us and near objects seem to speed backward in the direction opposite to our movement.

5. *Gestalt Effects*. The figure-ground distinction (p. 67) and the "good-figure" concept are considered by Gestalt psychologists to be native or physiological factors in perception, though many psychologists believe them to be primarily psychological. According to Gestalt theory, the figure intrinsically has better form than the ground; it has, too, more character or "thingness" and hence tends to stand out from the ground. The figure-ground relationship has relevance in auditory as well as in visual perception. The figure in music is higher in pitch than the ground, is louder and has different timbre. Thus a melody "stands out" as figure against its accompaniment as ground.

The geometric illusions have long been of interest to psychologists because they furnish clues to the perception process. Geometric illusions belong in the category of normal false perceptions—false because they can be shown to be erroneous, and normal because every one gets the illusory effect to some degree. Figure 59 shows the famous Müller-Lyer illusion; it is representative of a great many diagrams of the same sort. The two horizontal lines are of the same length, but the upper seems longer. This illusion has been variously explained as owing to eye movements, to confusion of the lines with the angles, to the relativity of all

judgment, and in other ways. The Gestalt psychologist considers the geometric illusions to be instances of "good-figure" theory. A three-quarters circle tends to be seen as a complete circle, and a broken triangle as a complete triangle (see p. 68); that is, each figure expresses its essential characteristics as fully as it can—tends to be as good as possible. The upper line in Figure 59 is extended (and hence is seen as longer) and the lower contracted (and hence is seen as shorter) to fit the configuration of the out-turning and the in-turning angles.

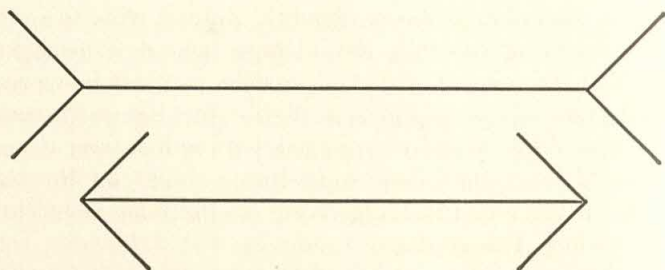


FIG. 59
THE MÜLLER-LYER ILLUSION

8. EXPERIENCE AND LEARNING IN PERCEPTION

If reference is made again to Figure 53, it will be noted that, with reference to the physical object, the retinal image is inverted. This inversion is due to the refractive mechanism of the lens. How is it that we see objects right side up when their retinal images are upside down? This question, which has often been proposed as a philosophical enigma, is puzzling only if we fail to appreciate properly the enormous rôle of experience and learning in perception. What we see is not the inverted retinal image, but the object in the field from which light-rays are reflected. The image on the retina is part of the total physiological process involved in getting the perception; but we are no more conscious of it than of the sensory impulse traversing the optic nerve. It must be remembered, too, that the retinal image is spatially correct with reference to its various parts, even though inverted, so that objects are perceived in their right relations. Up-ness and down-ness, right and left, what we see and how we see it depend upon all of the factors which we have mentioned plus hundreds of associated tactual and muscular experiences and contacts. We see things, feel them, fall over and walk to them, in many ways coming to know them.

The fact that objects do not always correspond to our sensory knowledge of them illustrates the phenomenon of *constancy*. Not only do objects re-

main constant as to size and form, but colors are also constant. A sheet of white paper is white, even in dull light, while coal is still black in bright sunlight.

An interesting experiment carried out by Stratton illustrates the phenomenon of object constancy. Stratton fastened before his right eye a tube containing two convex lenses. This instrument, of course, gave a retinal image which was right side up, and hence an inversion of objects appeared in the visual field. The left eye was covered over by a hood which excluded light, but allowed free ocular movements. Over a period of eight days (a total of eighty-seven hours), Stratton wore this instrument, meanwhile observing carefully any changes and new arrangements in the visual field. To prevent, as far as possible, conflicts between the new and the old (normal) perceptions, at night when the instrument was removed the eyes were carefully bandaged. On the first day, as might have been predicted, everything was topsy-turvy; things on the right were reached for on the left and things lying on the floor were reached for toward the ceiling. This confusion gradually abated, however, and by the end of the third day little was left of the nervous conflict so clearly apparent at first. By the eighth day the confusion in the visual field had almost disappeared; things could be reached for in their correct positions, and objects appeared normal (right side up) again. Stratton notes that even then there were sudden slips or inversions, which could usually be corrected by reaching for the object or moving toward it.

This experiment shows how the retinal image, although necessary to vision, is but one cue in the whole process of perceiving objects in space. It demonstrates, furthermore, how flexible are the connections between hand and eye and how readily modifiable and adaptable are our various sensory experiences. Ordinary everyday experiences confirm this. We learn to shave in a mirror, although right and left are reversed, and to operate a microscope, although the field is inverted. To repeat, the retina simply gives us the image in its proper relations; what we see depends upon the myriad associated tactual and muscular impressions called experience.

9. THE NATURE OF EYE MOVEMENTS

In describing briefly on pp. 302 to 306 the various factors which contribute to our knowledge of position in space, we did not deal directly with the rôle of eye movements in visual perception. It is readily recognized, however, that this must be a valuable factor, for our knowledge of objects and their structures is not obtained by steady fixation but by a series of short "views" as the eyes flit over the object and its surroundings. Careful studies by psychologists have shown that eye movements fall

into two main classes, "jump" movements and "pursuit" movements. Jump (or saccadic) movements occur when the eyes are moved voluntarily over an object; pursuit movements, when the eyes are controlled, so to speak, by the moving objects upon which they are fixated and move along with them.

Dodge has made probably the most careful studies of angular eye movements of the jump type, using the method of photography. Early investigators, for example Huey, had attempted to register eye movements by mechanical means. A small disk was glued to the cornea of the subject's eye, and a fine thread led from this to a delicate writing point. When the eye moved, its excursion was recorded by the writing point on a smoked drum. Besides obvious objections to this procedure from the standpoint of the subject, there was inaccuracy and lag in the recording by this system. The tracings, too, tended to blur at the beginning and end of the movement, thus making measurement unreliable. Photography of the eyes in motion eliminates most of these difficulties. With his subject seated in a dark room, and position of the head kept constant by means of a head rest, Dodge reflected a beam of light from the cornea upon a moving sensitive plate. As the plate moved downward vertically between two grooves, the subject was instructed to move his eyes from a given point to another in the same horizontal line. Focus, or rest points, showed up as bright spots on the film, intermediate movements as white lines, or streaks.

The time required to shift the focus from point *A* to point *B* was measured by photographing an oscillating spring pendulum along the edge of the moving plate. The vibrations of this time marker showed up as a series of toothlike indentations which could easily be counted. Since the pendulum rate of vibration was known, it was possible to measure the time intervening between the fixation at *A* and then at *B*. As might be expected, the time of an angular movement of the eyes depends largely upon the extent of the excursion. Dodge found that the time required for the eyes to move through an angle of five degrees is about .03 second, the time for twenty degrees is about .06 second, and the time for forty degrees is .1 second. These results are based upon the records of three subjects. The reaction time of the eye to a beam of light cast upon the retina—i.e., the time intervening between the stimulus and the eye's movement—was determined by Dodge to be around .165 second.

The jerky character of eye movements in ordinary seeing has been clearly shown by Stratton, who also employed the photographic method. Contrary to casual opinion, the eye does not move in a smooth sweep, unless the movement is of the pursuit variety. In such movements, as previously noted, the eye simply follows the moving object, and is not

actively controlled by the observer. In looking at an object or a picture Stratton found that the eye moves in a series of jerks or jumps, even though the subject thinks he is moving his eyes smoothly. Like Dodge, Stratton placed his subject in a dark room and photographed the eye movements by reflecting light from the cornea upon a sensitive plate. The way in which the eye moves in looking around a circle is shown in Figure 60. The dots represent pauses.

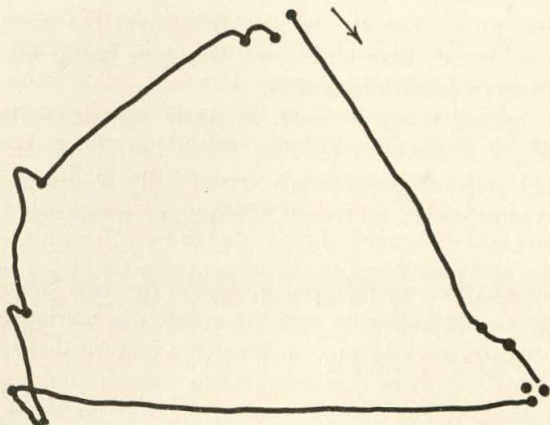


FIG. 60
PHOTOGRAPHIC RECORD OF THE EYE'S MOVEMENT IN TRACING THE OUTLINE
OF A CIRCLE
(After Stratton.)

10. EYE MOVEMENTS AND OTHER FACTORS IN READING

One of the first studies of eye movements in reading was carried out by Dearborn, who used the method of photographic registration employed by Dodge and Stratton. Dearborn found that the number of fixations per line for ordinary printed matter varied from four to seven. The number of "stops" in a line is variable, depending upon the difficulty of the material as well as upon the education and mental ability of the reader. Illumination and print size are also important. Tinker has corroborated Dearborn's finding that the *fixation pauses* of the eye in reading consume 90 to 95 per cent of the total reading time. The utilization of so much of the total reading time in fixation, and the high speed of "jump" movements, make clear perception during eye movement practically impossible. In spite of this our visual world is not a patchwork of clear visual spots, filled in between by blurs, but is continuous and spread out before us.

There are many reasons why this is true. In the first place, the visual field, when the eyes are at rest, encompasses much more than the par-

ticular object focused upon; there is considerable "overlap" or continuity from one fixation pause to another. There is clear vision, too, of moving objects in pursuit movements, although in following objects which are rapidly moving the background is often blurred. More important than purely visual factors is the continuous "filling-in" by data got from the other senses. This background of information in terms of which our immediate perceptions are "sized up" is sometimes called the "apperceptive mass." A familiar example is the speed with which educated adults read easy or familiar prose. The eye sweeps along over the page, hitting important or "key" words here and there, the gaps being filled in by the context and by associated information.

Fatigue in reading is not as great as many students believe. In fact, careful studies of continuous reading (both hard and easy materials) by high school and college students reveal little measurable fatigue as shown by eye movements, increased blinking, or weakened heart-rate.

Early in the chapter we referred briefly to the two theories of visual space perception, the nativistic and the empiristic. Perhaps most of the facts here presented have seemed to favor the empirical theory, although the physiological factors of accommodation, convergence, and muscular strain weigh heavily. No one, of course, can say definitely just where the influence of native factors leaves off and the influence of learned factors begins. Learning could do nothing without the native equipment of a spatially correct retinal image and the physiological factors which furnish cues to movement. Nor could the eye alone build up visual space as we know it unless aided by associated observations from the other senses. Both sets of factors are important and necessary, and we could dispense with neither.

BIBLIOGRAPHY

1. Carmichael, L., & Dearborn, W. F. *Reading and visual fatigue*. Boston: Houghton Mifflin Company, 1947.
2. Dearborn, W. F. The psychology of reading: an experimental study of the reading pauses and movements of the eye. *Arch. Phil. Psychol. & Sci. Method*, 1906, 4.
3. Dodge, R., & Cline, T. S. The angle velocity of eye movements. *Psychol. Rev.*, 1901, 8, 145-157.
4. Galambos, R., & Davis, H. The response of single auditory-nerve fibers to acoustic stimulation. *J. Neurophysiol.*, 1943, 6, 39-57.
5. Graham, C. H., and Hartline, H. K. The response of single visual sense cells to lights of different wave lengths. *J. gen. Physiology*, 1935, 18, 917-931.
6. Huey, E. B. On the psychology and physiology of reading. *Amer. J. Psychol.*, 1900, 11, 283-302; 12, 292-312.

7. Morgan, C. T., & Steller, E. *Physiological psychology* (2nd ed.) New York: McGraw-Hill Book Company, Inc., 1950. Chapters 9 and 10.
8. Myers, C. S. *A textbook of experimental psychology*. New York: Longmans, Green & Company, 1925, chapters 3, 4, 6, 7.
9. Seashore, C. E. *The psychology of musical talent*. New York: Silver, Burdett and Company, 1919.
10. Starch, D. Perimetry of the localization of sound. *Psychol. Monog.*, 1905, 6, No. 28, 1-45; 1908, 9, No. 38, 1-55.
11. Stevens, S. S., & Newman, E. B. The localization of actual sources of sound. *Amer. J. Psychol.*, 1936, 48, 297-306.
12. Stratton, G. M. Vision without inversion of the retinal image. *Psychol. Rev.*, 1897, 4, 341-360, 463-481.
13. Stratton, G. M. Symmetry, linear illusions and the movements of the eye. *Psychol. Rev.*, 1906, 13, 82-96.
14. Tinker, M. A. Time taken by eye movements in reading. *J. genet. Psychol.*, 1936, 48, 468-471.
15. Tinker, M. A. Readability of book print and newsprint in terms of blink-rate. *J. educ. Psychol.*, 1948, 39, 35-39.
16. Wever, E. G., & Bray, C. W. The nature of acoustic response: the relation between sound frequency and frequency of impulses in the auditory nerve. *J. exp. Psychol.*, 1930, 13, 373-387.
17. Wever, E. G., & Bray, C. W. Present possibilities for auditory theory. *Psychol. Rev.*, 1930, 37, 365-380.
18. Woodworth, R. S. *Experimental psychology*. New York: Henry Holt and Company, 1938, chapters 21, 22, 23, 26.
19. Young, P. T. Auditory localization with acoustical transposition of the ears. *J. exp. Psychol.*, 1928, 11, 399-429.

15

CATTELL'S STUDIES OF REACTION TIME

1. REACTION TIME AND PSYCHOLOGICAL RESEARCH

ONE OF the most direct ways in which the processes of perception, discrimination and choice may be subjected to quantitative study is to measure the *time* it takes an individual to perceive and report upon the likenesses or differences among several stimuli. Various experiments in the psychological laboratory have been concerned with the measurement of time of response, or *reaction time*, i.e., with the accurate recording of the time intervening between the application of a stimulus, the sound of a buzzer, say, and the motor response of the subject, e.g., lifting the finger from an electric key. Hope that the study of reaction time would provide a sort of mental chronometry whereby the duration of such processes as judgment and will could be determined has not been realized. But measurement of reaction time has proved to be valuable in a variety of ways.

The primary interest of many experimenters has been in the establishment of reaction time *norms* for various kinds of stimuli, viz., light, sounds, and touches; and in studying individual differences in reaction to such stimuli. These measurements have been useful in the selection of men for jobs requiring speed and accuracy of response. Other workers have studied the effect of different stimulus conditions—the subject meanwhile being kept relatively “constant”—by changing the quality, intensity, or duration of the stimulus. Or different parts of the body have been stimulated and, in the case of light, say, the results compared when the retina is excited at the fovea, and at varying distances from this central point. Still other researchers have varied the *condition of the subject*, meanwhile keeping the task relatively constant. In this way it is possible to measure the effects of drugs, fatigue, practice, incentive and punishment, and other factors upon the reaction time of the individual.

In America much of the interest in reaction time in psychology may be traced to the work and influence of James McKeen Cattell. Cattell was one of the first American students to take his doctor's degree in Wundt's laboratory at the University of Leipzig. Wundt's laboratory was established

in 1879, and was the first psychological laboratory in the world. Cattell's dissertation was printed in English in 1886 under the title *The Time Taken Up by Cerebral Operations*, and was followed by other important reaction time studies. During his thirty years as professor of psychology at Columbia University, Cattell and his students published many studies of the time relations of mental processes. Some of these studies will be described in this chapter. Suffice it for the present, to consider the influences which prepared the way for the interest of psychologists in the time measurements of mental phenomena.

2. THE HISTORY OF THE REACTION TIME EXPERIMENT

The reaction time experiment has a long and interesting history. A story which dates back to 1796 tells of an unlucky assistant in the astronomical observatory at Greenwich, England, who lost his job because he persistently recorded the instant when a star passed across the meridian with too large an error—his reaction time or “personal equation” was abnormally long. The first publication of the results of comparative tests showing fairly large and persistent individual differences in recording the transits of stars was that of the German astronomer, Bessel, in 1822; too late, unfortunately, to save the slow-reacting assistant mentioned above. In 1850 Helmholtz, the great physiologist, used the reaction time method to measure the speed of nerve conduction. Working with the motor nerve of the frog, Helmholtz stimulated a nerve attached to a muscle at a point some distance from the reacting muscle. The time intervening between the application of this stimulus and the beginning of movement in the muscle (often called the *latent* period) was recorded, as was also the shorter time-interval elapsing between the application of a stimulus nearer the muscle and the appearance of movement. The difference between these two reaction times gives the time taken by the impulse in traveling along the nerve between the two points of stimulation. For the frog, Helmholtz reported the speed of motor nerve conduction to be 27 meters (about 89 feet) per second; in man he calculated the speed of motor nerve conduction by the same method to be about twice as great.

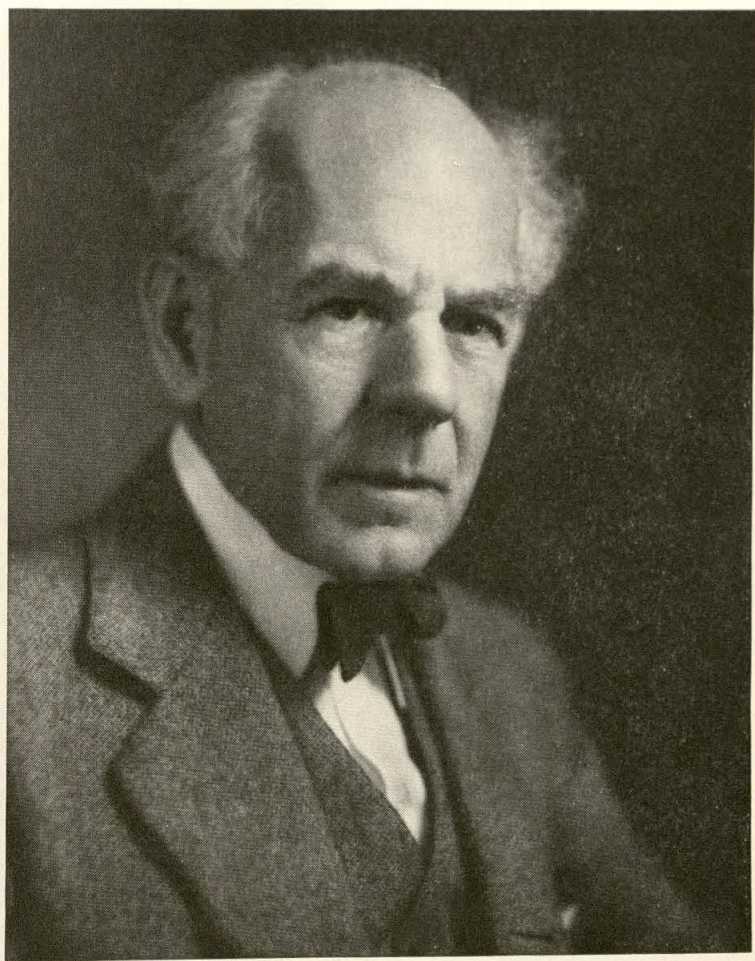
Helmholtz also applied his method to the measurement of sensory nerve conduction. He assumed that if two points on the same sensory nerve, one, say, on the foot, and one on the upper leg, were stimulated in successive experiments, the subject's response in both cases being the same, the difference in reaction time would be a measure of the time taken by the impulse in traversing the nerve length between the two stimulated points. Unfortunately, the results of these experiments were so variable that Helmholtz's determinations were highly uncertain. Probably the main

reasons for this variability are the many influences, both facilitating and inhibiting, which the impulse must undergo in passing through the nerve centers in the brain and spinal cord.

The many investigations which followed those of Helmholtz gave results for the most part widely divergent from his, and also from each other. Finally, in 1895 Cattell and Dolley published the most thorough investigation into the problem of nervous conduction in man by the reaction time method. Cattell and Dolley applied electrical stimuli at two points on the median nerve of the arm 30 centimeters apart, and also at two points on the posterior tibial nerve of the leg 50 centimeters apart. The reaction movement was the same for all stimuli. Large variations were found in the velocities of the sensory impulse as so measured, the times being 31 meters per second for one observer, and 65 meters per second for another. Considerable variability, as well, was encountered in each observer. These results led Cattell and Dolley to suggest that the differences in reaction time found must be due, not to differences in the speed of the nervous impulse, but to the variability of the central connections involved, as well as to qualitative differences in the sensations aroused at the different points of stimulation. This explanation was substantiated by the interesting finding that either hand will react more quickly when it, rather than the other hand, is stimulated. According to these authors, this obviously means that some sensori-motor connections, owing possibly to innate linkages and more probably to practice, are more closely knit than others, so that the reaction time varies even when the *same* length of nerve fiber is traversed. Because of complications in the central connections, Cattell and Dolley concluded that the velocity of the nervous impulse cannot be adequately measured by reaction time methods. It should be added in ending this section that physiologists using other and improved methods have determined the speed of the nerve impulse in man to be around 70 meters (about 225 feet) per second. This is much faster than the times found by the early investigators who used reaction time methods. The speed of the sensory nerve process is probably about the same as that of the motor.

3. EARLY STUDIES OF REACTION TIME

The first studies of the psychological factors influencing simple as well as complex reactions were those of the Dutch scientists Donders and De Jaeger in 1865. In the *simple reaction* experiment as now set up in the laboratory, the subject knows beforehand what stimulus he will receive, as for example, a flash of light or a sound. Usually his instructions are to make a single prescribed movement, such as lifting the finger from



JAMES MCKEEN CATTELL

1860-1944

an electric telegraph key (which is being held down) as quickly as possible when the stimulus is given. After practice, the simple reaction becomes highly automatic and almost reflexive. In the *discrimination reaction* experiment, as devised by Donders, two stimuli (e.g., two notes of different pitch or two differently colored lights) are employed, and the subject is instructed to react to the one and remain quiet to the other. When, in addition to the two stimuli, the subject is given *two* reaction keys, and is instructed to react with the *right hand* if one stimulus appears and with the *left* if the other stimulus appears, we have what is called a *choice reaction*.

Donders, as well as other early workers, believed that by *subtracting* the simple from the discriminative reaction time, "discrimination times" could be obtained. Donders also held that by subtracting discrimination reaction time from discrimination *plus* choice, "will" or "choice" times could be found. This procedure is called *elimination by subtraction*. It assumes that a complex response is in reality a simple response plus certain added central processes, and that the time taken by these processes may be determined by subtracting the time of the simple reaction.

Cattell, as early as 1885, criticized the view that a discriminative reaction may be analyzed into a simple reaction plus some more or less constant "will time" or "perception time." Subtracted times, he held, were simply measures of the increased complexity of the total process, or of the increased difficulty of the task. Other investigators who questioned elimination by subtraction as a method were Erdmann and Dodge (1898), and Ach (1905). The latter, especially, argued that the different preparation of the subject for a simple and for a discriminative reaction made the two very different psychologically. When faced by several stimuli, or given the choice of several responses, the *situation*, said Ach, is very different from that in the simple reaction time experiment. As a result a different attitude is set up in the reactor, so that the whole process is different, not merely the *central* part of it. Moreover, as was urged by other experimenters, there is no introspective evidence that, when one mental process is added to another, the first retains its identity unchanged.

Cattell's view of what happens in reaction time experiments is very different from that of Donders, and also from the explanation advanced by Wundt. At the ready signal, he said, the subject, in making a simple reaction, concentrates his attention upon the stimulus which is to appear, e.g., a red light, or an electric shock, and upon the finger which is to react. Hence, the nervous pathways between the eye or skin→brain→finger are especially well prepared or open when the stimulus comes. In the discriminative or choice experiment, obviously more "switches" and more pathways must be "set" because of the increased complexity

of the situation. This introduces an element of suspense or indecision which in turn necessitates greater preparation and makes for longer reaction time. Wundt had regarded the whole reaction process as analyzable psychologically into (a) perception of the stimulus, (b) apperception (Wundt's term for the recognition or comprehension of the stimulus), and (c) will, or the release of the impulse to react. This analysis, said Cattell, is not psychologically accurate. Apperception, or the clear recognition of the stimulus, comes if at all *after* the reaction has been made, while will exists simply in the preparation for movement and becomes less and less a factor as the reactions grow more and more automatic.

In justice to Wundt it should be said that, while his analysis probably failed to describe the somewhat simplified laboratory experiments on reaction time, it does fit quite well many unprepared reactions of everyday life. Suppose, for instance, that a child runs in front of an oncoming bus, and that the driver applies his brakes just in time to prevent an accident. Here, in Wundt's terms we have perception of the child, apperception of the danger in the situation, and will, or the release of impulses which stop the bus. But these psychological processes cannot be tagged in any such clear-cut fashion in the simple laboratory situation of lifting a finger from an electric key as soon as one hears a sound or sees a light. Nor do they appear as separate and distinct processes when reactions become habitual and automatic as they are for the experienced bus driver.

4. APPARATUS USED IN MEASURING REACTION TIME

We have said that the simple reaction time experiment measures the time-interval between a single and oft-repeated stimulus and a prescribed and constant response. In the usual laboratory setup, the subject sits with his forefinger depressing an electric key; and at the appearance of the stimulus he reacts by lifting his finger from the key as quickly as possible. Various methods have been employed in the laboratory for measuring the time-intervals between stimulus and response. All of these make use of a chronoscope or some other time-measuring device. The Hipp chronoscope, an instrument which measures in units of .001 second,¹ has perhaps been most widely used in the past by psychologists (see Figure 61). Apparatus for checking the accuracy of the Hipp chronoscope, as well as many improvements in the instrument itself, were made by Cattell and his students. The Bergstrom chronoscope, and the Dunlap chrono-

¹ .001 sec. has usually been represented by 1σ in psychological literature. Thus, 100 σ and .1 sec. are the same. Unfortunately, the sign σ is a bad choice, because of its widespread use in statistics to denote the standard deviation. A better time-unit is milliseconds, or ms.; 1 ms. = .001 sec., and 100 ms. = .100 sec.

scope, are also widely used. In addition to instrumental measurements, graphic records of reaction time may be obtained by using an electrically driven tuning-fork and a smoked drum. A tuning-fork which vibrates 100 times a second, say, is set to mark off a time-line on a revolving drum. Beneath this time-line the appearance of the stimulus and the subject's

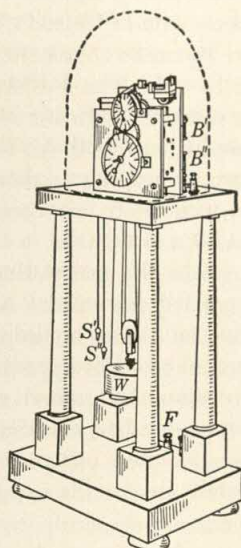


FIG. 61

THE HIPPI CHRONOSCOPE

The clockwork of this chronoscope is driven by the weight, *W*, which is raised by a key fitting in the center of the lower dial. The chronoscope is started by pulling *S'* and stopped by pulling *S''*. The two recording dials are divided into hundredth parts. The hand on the upper dial revolves ten times every second, each division on the dial corresponding to a thousandth part of a second. The hand on the lower dial revolves once every ten seconds, each division on the dial corresponding to one tenth of a second. Reaction time is calculated from both dials, the units and tens from the lower and the hundredths and thousandths from the upper dial. The speed of the dial hands is controlled by a small steel tongue which is accurately tuned to vibrate 1,000 times per second. *B'* and *B''* are electromagnets which control the starting and stopping of the dial hands. *F* is a binding-post. Various control instruments have been devised to check the accuracy of the readings given by this chronoscope.

response are recorded by means of electric markers. The graphic method is dependable and is often used as a check on other methods; but it is too slow for most practical work.

All of the devices described above may be used in the measurement of complex as well as simple reactions. Experiments in complex reaction time may be set up either through the addition of extra stimuli, or through the requirement of more than one response. In the Donders discrimina-

tion reaction experiment previously described, either of two stimuli was presented, to only one of which the subject was instructed to react. The obvious disadvantage of this setup is that records are obtained in only about one half of the trials. In Donders' discrimination and choice experiment, two stimuli were employed and two responses, the subject reacting with the right hand if the one stimulus appeared, and with the left hand if the other appeared. Henmon, one of Cattell's early students, simplified and improved this procedure by an arrangement in which the two stimuli were always presented side by side. The subject was instructed to react to a previously designated stimulus with his *right* hand if it appeared to the right of the other stimulus, and with his *left* hand if it appeared to the left of the other stimulus. As will be readily surmised, reaction time is considerably slower when conditions are complicated than when only one stimulus and one response are utilized.

Reaction time is still longer in the *associative* type of response. Here the stimuli may be words (or lists of words) to which the subject is to react with the first word which comes to mind (free association); or, the subject may be instructed to give the opposites of the stimulus words, or responses bearing some other designated relation, e.g., part-whole, adjective-noun, verb-object (controlled association). The color-naming and form-naming tests illustrate other varieties of controlled association. The usual method in associative reaction is to take the total time required by the subject to give the correct responses to a list of words and then determine his average reaction time for each word. The one-fifth second stop-watch is probably accurate enough for reactions of this kind; for finer work, a one-hundredth second electric stop-clock may be employed.

5. FACTORS INFLUENCING SIMPLE REACTION TIME

Many conditions influence the speed of reaction besides the complexity of the stimulus-response situation. Most of these factors have been studied at one time or another by Cattell and his students, as well as by many others, and some of the more important findings will be considered in this section.

1. *Sense Organ.* Reaction time varies with the sense organ stimulated. Cattell reported the simple reaction time to light to be about 150 ms., with some variability (8-10 ms.), and the simple reaction time to sound to be about 120 ms. Simple reaction times to stimuli applied to the different sense organs, as determined by many investigators, may be summarized as follows:

<i>Type of Stimulation</i>	<i>Reaction Time</i>
Visual	150 ms.— 200 ms.
Auditory	120 - 160
Tactual	110 - 160
Olfactory	200 - 800
Gustatory	300 -1000
Pain	400 -1000
Temperature:	
Cold about	150 ms.
Warmth about	180

Probably the slow reaction time to taste and smell is owing to the chemical nature of these senses and to the difficulty in getting stimuli directly to them. The "free nerve endings" of the pain sense are fairly close to the surface, but are relatively slow to react. The fact that hot and cold stimuli must act by conducting heat or cold into the skin to the sensory end organs will obviously slow up the reaction time to such stimuli.

The part of the sense organ stimulated makes a difference in reaction time. Poffenberger, a student of Cattell, found that when the retina is stimulated with a light beam at points to the right and left of the fovea, reaction time is longer than it is when the fovea itself is excited. Moreover, reaction time becomes longer and longer as the stimulus is moved out from the central point to the periphery of the retina. Reaction time is faster to stimuli applied to the hand than to stimuli applied to the forehead or foot, as a result probably of closer nerve center connections in the first instance.

2. *Intensity, Size, and Duration of the Stimulus.* Cattell studied the influence of all three of these factors upon reaction time as early as 1886. More accurate work was done later by Froeborg, under Cattell's direction. Working with graded light intensities (light reflected from grey and white paper), Froeborg found that reaction time increased, but very slowly, as the intensity of the light decreased. When reaction time to a given medium intensity, for example, was 191 ms., the reaction to a stimulus one half as bright was increased by only 3 ms.; while the reaction to a stimulus one tenth as bright was increased by 17 ms. The same general result was found when the *size* (area of a square surface reflecting light) and the *duration* of the stimuli were decreased. The reaction time increased in both cases, but somewhat more slowly for duration changes than for size changes. Steinman, using better methods than Froeborg, found also that reaction time increases with intensity but only up to a medium intensity level. Beyond this point, greater intensities gave slower reaction times. Froeborg found that reaction time to sound increased fairly rapidly when the intensity of the sound was steadily diminished.

3. *The Foreperiod and the Reacting Movement.* The foreperiod, that is, the time-interval between the signal "Ready" and the presentation of the stimulus, is quite important in reaction time work. During the foreperiod, the subject "gets set," so to speak, to react to the expected stimulus as quickly as possible. If the foreperiod is short, i.e., less than 1 second, the subject may be caught before he is entirely ready, and in consequence react more slowly than usual; if the foreperiod is long, say 10 seconds or more, the subject is likely to lose his "edge" and hence react too slowly. Cattell placed the optimum foreperiod close to 1 second. This interval, he said, is about as long as the nerve centers can be kept in a state of preparedness; longer times (up to 15 seconds) were found to increase the reaction time as well as the variability of response. Other investigators placed the best foreperiod at from 2 to 3 seconds. Breitwieser, working in Cattell's laboratory with improved technique and more subjects, found the optimum foreperiod to be between 2 and 4 seconds, with, however, some individual differences. This determination has been verified by Woodrow, who also found 2 seconds to be the most favorable foreperiod.

When stimuli are given at regular intervals, the reactor builds up an expectancy or "set" and his reaction time becomes shorter. Intervals longer or shorter than those expected will now lead to an increase in reaction time. This was shown in an experiment (Mowrer) in which a regular stimulus interval of 12 seconds was first established. Later when intervals from 3 to 24 seconds were introduced irregularly, the reaction time was considerably increased.

The influence of the reacting movement upon time of response has already been touched upon. Either hand, for instance, responds more quickly to a stimulus applied to it than to one applied to the other hand. It has been shown, too, that releasing rather than depressing the reaction key gives a less variable response. In a careful study of the movements of the hand in reacting to a stimulus, carried out in the Yale laboratory in 1905 by Judd, McAllister, and Steele, it was discovered that the hand does not maintain a steady pressure on the reacting key during the foreperiod. On the contrary, there is a fluctuation from stronger to weaker pressure, the key sometimes being completely released to give a "premature" reaction. Whenever degree of readiness to react has reached a high point, the response may easily be premature, as when a runner springs from his mark before the signal is given. The Yale experimenters found, too, that often the first reactions of unpracticed subjects show a quick depression of the key before it is released. This counter-movement takes time and consequently slows up the reaction. With practice it tends to be inhibited in favor of the correct release movement.

4. *Practice, Attention, Distraction, and Fatigue.* Practice, according to Cattell, has little effect upon time of response after the first few trials—in which the subject's reactions often vary markedly. The effect of distractions, such as the beating of a metronome, or the performance of mental addition during the experiment are also relatively slight provided the subjects are highly practiced. For the first type of distraction the reaction time may increase from 2 to 10 ms., for the second from about 20 to 30 ms. Woodrow has investigated with great thoroughness the rôle of attention and practice in reaction time. He found that reaction time varies considerably with the amount of attention given the task. Degree of attention was measured by the prolongation in reaction time brought about by introducing foreperiods of varying lengths. For stimuli of "moderate" intensity, the subject's attention was less affected for touch than for sound or light. Evans, using flashing lights, noises, and touch stimuli as distractions, found that these extraneous factors increased the reaction time markedly at first, but that their influence was greatly lessened with training. Both trained and untrained observers were affected by distractions, however, the greatest disturbance occurring when both the distraction and the stimulus to which the subject was instructed to respond affected the *same* sense organ. Cassel and Dallenbach also report that distractions tend to lengthen reaction time, the increase for two observers ranging from 3 ms. to 37 ms. Fatigue, strangely enough, has an almost negligible effect upon reaction time. In his own case, Cattell found his reaction time to decrease very slightly even after a day of continuous reacting.

5. *Incentives, Punishment, Drugs, and Age.* The effect of incentive—encouragement or mild praise—upon reaction time is to increase it about 8 ms. Negative incentive, or punishment, as for example giving the subject an electric shock if he fails to react within a given time, speeds up the reaction time 20 ms. or more. Drugs have a variable effect upon reaction time. Coffee and tea appear to shorten it; small doses of alcohol first shorten and then lengthen it; morphine, ether, and chloroform usually lengthen it. As might be surmised, response time is slower and more variable in childhood and in old age.

6. THE EFFECT OF ATTITUDE UPON REACTION TIME

In addition to the factors discussed in the last section, there is still another determinant in reaction time experiments which has excited much controversy, and which is important enough to be considered at some length. This has to do with the *direction* of the subject's attention, or preparatory set during the foreperiod. Several early workers on reaction time

in Germany, notably Lange, in 1888, observed that when the subject in a simple reaction time experiment directed his attention specifically toward the awaited stimulus, his time of response was considerably *longer* than when he fixed his attention upon the movement to be made. Lange called the first kind of reaction "sensorial," and the second "muscular," or "motor." The first type of response, he said, is always longer than the second, the difference being as much as 100 ms. The distinction between sensorial responses and motor responses was accepted by most contemporaries of Lange as valid, although the differences found were not always as great as those he had reported.

Lange's Attention Theory, as this view was called, was challenged by Baldwin's (1895) and Flournoy's (1896) Type Theory. These two workers found among their subjects some whose motor reactions were faster than their sensory reactions and others who gave quicker sensory responses than motor responses. They suggested, accordingly, that individuals are natively sensory, motor, or indifferent in type, and that each type reacts at his best when allowed to follow his natural inclination. Thus, while accepting the distinction between sensory and motor responses, Baldwin and Flournoy reached a very different conclusion with regard to the differences in speed between the two types from that of Lange and his group.

Still a third explanation of the difference between sensory and motor responses has been given in terms of practice and habituation. This was the view advanced by Cattell in 1892 and by Angell and Moore in 1896. Cattell had found no reliable differences between the sensory and motor types of reaction in the case of practiced observers. Whether a new subject is sensory or motor at the start, he said, is largely a matter of the instructions given, accidental direction of attention, and previous training; after practice, there is little or no difference between the two attitudes. Later work, done by Breitwieser, under Cattell's direction, tended to confirm this view. Breitwieser first took a long series of reactions in some of which his subjects were verbally instructed to give sensory, and in others motor, responses. Following these experiments, Breitwieser forced the attitude of his subjects into the sensory or motor category by means of the following ingenious arrangement. To insure a sensory attitude, he presented sometimes the one and sometimes the other of two closely similar stimuli, and required his subjects, after reacting, to report which of the two had been given. This necessitated close attention to the stimulus. To make sure that the reaction was of the motor type, Breitwieser used a reaction key which varied somewhat in the resistance which it offered to the subject's reacting finger. The subject was instructed, after responding, to tell which of the two resistances (the greater or less) had been employed. Here, clearly enough, it was the reacting movement which

needed special attention. Finally, in order to guarantee a neutral, or indifferent attitude, the subject was instructed to identify after each reaction *both* the key pressure *and* the stimulus presented.

Breitwieser's results indicated in general a very slight advantage—averaging about 20 ms.—for the verbally instructed motor over the verbally instructed sensory reactions. He found little difference between those sensory responses in which the subject was verbally instructed to attend to the stimulus, or forced to attend. However, the verbally instructed motor responses were much faster (in fact, the fastest of all) than those motor reactions in which the subject was required to report upon the key pressure. When the subject was instructed to report *both* upon key pressure and stimulus difference, his reaction time was the slowest of all, the time being almost as great as that in a discriminative reaction. The upshot of this whole controversy would seem to be (1) that there is a real distinction between the sensory and motor attitudes; (2) that untrained subjects are probably for the most part sensory at the start, becoming motor as the novelty wears off and the need for close attention to the stimulus becomes less; and (3) that the motor, being the more highly practiced attitude, is the faster type of response.

7. ASSOCIATIVE REACTIONS

While Cattell was interested in the study and analysis of the factors influencing simple reaction time, he was also interested in the more elaborate discriminative and associative reactions. He found, for example, wide variations in the time it takes persons to multiply two numbers presented at the same time, and that a subject can name objects or pictures in his own language more readily than in a foreign tongue. All this seems fairly obvious, and is probably no more than one might expect. But many other findings are by no means so evident. Thus, Cattell found that when a word indicating a *part* of an object is the stimulus, e.g., *roof*, it takes longer to give the name of the *whole* object, e.g., *house*, than to give the part when the whole is presented, e.g., *pencil—point*. Also, it is easier to go from a special to a more general class, e.g., *cat—animal*, than in the reverse direction. Opposites are given with the quickest reaction times, but these times vary greatly, the differences here, as in the other cases, being due largely to familiarity, training, and frequency of usage.

In other experiments with words and language, Cattell found that while the reaction time to a single word exposed alone is about 360 ms., this time is reduced to 200 ms. per word when a series of words is exposed together. Moreover, a higher speed of reading per word is obtained when two words instead of one are presented, three instead of two, and so

on up to four or five. These results, taken together with experiments made with the tachistoscope (in which it had been demonstrated that one can group as many as two or three words together in an instant of exposure) indicate quite clearly the great importance of *overlapping* in efficient verbal response. They show, too, that an individual's reaction to a phrase or sentence composed of several words is not simply the sum total of his time reactions to the *separate* words, but is rather the response to a group of words taken as a *larger unit*. This suggests that growth in reading ability, as well as in many other highly skilled performances, must be due to the fact that reactions are made as wholes to larger and larger groups of discrete impressions. Responses in larger units enter into the performance of the highly trained musician playing from score. For his experiments with language, Cattell devised a lip key and a voice key, both of which were distinct improvements on similar devices used up until then in recording oral responses.

Cattell's experiments have had an interesting effect upon methods of teaching reading. On the theory that learning should proceed from the simple to the complex, children for years had been laboriously taught their *a, b, c's*. These elements learned, pupils were then taught how to put letters into words, words into phrases, and phrases into sentences. This is not the way in which language is learned naturally. In fact, it is a logical rather than a psychological method of learning. Cattell's experiments showed that words are perceived and reacted to as *wholes*, not as clusters of letters; that one fails to see errors in a printed page because he is not looking for letters but reading words. In many modern schools children are now being taught to read words from the start and letters are learned afterward. This more natural method is often employed, too, with adults who wish to acquire a rapid reading knowledge of a foreign language. The utilization of reaction time experiments in pedagogy is a good illustration of how a somewhat restricted laboratory result may have widespread practical outcomes.

8. THE DISCRIMINATION TIME METHOD

The last series of experiments which we shall describe is that in which a new psychophysical method was employed. The "discrimination time method" was reported by Cattell in 1902, and is based upon the assumption that differences in sensation are equal when it takes the same time to perceive them. In general, the smaller the difference, the longer is the time necessary to perceive and react to it. Response differences become, then, quantitative indices of stimulus differences or difficulty.

The most important study by the discrimination time method is that

of Henmon. Henmon's method was to present two stimuli simultaneously. Two keys were used for response, the subject being instructed to react to one of the two stimuli (which one was previously designated) with the right hand if it appeared on the right, and with the left hand if it appeared on the left. For example, suppose red and blue are the two stimuli, and the subject has been instructed to react to red. Under these conditions, when the stimuli are presented the subject must react with his right hand when the red is to the right of the blue, and with his left hand when the red is to the left of the blue. Some of Henmon's results with colors appear below.

<i>Colors to Be Discriminated</i>	<i>Discrimination Time</i>
White and black	197 ms.
Red and green	203
Red and blue	212
Red and yellow	217
Red and orange (mixed with 25% red)	252
Red and orange (mixed with 75% red)	271

It is clear that the smaller the *qualitative* differences between the two colors, the longer the reaction time. Using small differences in pitch Henmon obtained much the same results:

<i>Differences in Pitch (in Vibrations)</i>	<i>Reaction Time</i>
16	290 ms.
12	299
8	311
4	334

When the subject was instructed to react to the *shorter* of two lines, shown side by side, the following reaction times were obtained:

<i>Differences in length of line (in mms.)</i>	<i>Reaction Time</i>
From 10 to 13	296 ms.
10 12	305
10 11½	313
10 11	334
10 10½	345

These results show that the smaller the difference between the two stimuli, whether in vibration rate or in millimeters, the longer the reaction time to the difference.

Henmon's findings may be interpreted to mean (1) that the mental processes of apprehension and discrimination need more elaboration—and hence more time—when differences are decreased; and (2) that the

preliminary adjustment for the movement is less and less adequate as the difference between the two stimuli becomes smaller and smaller. Both of these conditions serve to explain the lengthened reaction time. If we are justified in assuming that perception differences are equal when the times required to discriminate them are equal, it appears from Henmon's results that tones 8 vibrations apart differ *perceptually* to about the same degree as lines differing by $1\frac{1}{2}$ millimeters (10 and $11\frac{1}{2}$).

9. SOME PRACTICAL APPLICATIONS OF REACTION TIME STUDIES

It should now be evident that the study of reaction time has passed through many phases and has excited many and various interests. Reaction time is measured today not so much for the purpose of analyzing its determinants or of finding "choice" or "will" times, but rather as a means of attacking practical problems of behavior. For example, the reaction time experiment may be employed (1) to measure the *difficulty* of a task in an objective way, or (2) as an index of an individual's *efficiency* under different conditions. Furthermore, reaction time may be taken as a means of *comparing* (3) individuals under the *same* conditions, or (4) the same individual under *different* conditions—the establishment of individual differences may be the primary interest. Again, the speed of a person's simple or discriminative responses may be compared with his quickness in solving problems or in performing other intellectual tasks to see whether a given individual may be characterized as generally quick or generally slow intellectually. In one study of this kind (Lemmon) memory for words, numbers and the like, was found to be related to speed of simple reaction, while the learning of verbal and other abstract relations was correlated with quickness of discriminative reaction. Few persons, however, can be fairly described as possessing a uniform speed-level at which they habitually work. Differences in familiarity with the tasks to be done, degree of previous practice, interest, and incentive play too large a rôle. A bookkeeper, for instance, who is ordinarily very slow may, after long practice, or with sufficient incentive, surpass in mental calculation—addition, say—an individual much faster in other respects.

In vocational selection, as well as in the diagnosis of aptitudes, the reaction time of the individual is often an important factor. It is apparent, for instance, that knowledge of the time required by an automobile driver or an engineer to apply his brakes at a given signal, or to react quickly in an emergency is of the greatest practical value. Ingenious devices for testing the reaction time of prospective drivers of automobiles or of street-cars have been devised by psychologists. Wechsler, for instance, constructed a dummy automobile, with steering-wheel, brakes, clutch, and

so on, which he used in testing prospective taxi drivers. The candidate was instructed to depress his clutch and apply his brakes at the flash of a yellow light on the board placed before him; at the appearance of other colored lights to react in various ways with appropriate hand and foot movements. Time of reaction and errors (wrong responses) were carefully recorded. After actual road experience it was found that, while men with the slowest reaction times had the most accidents, the fastest reacting men also had a large number of accidents. The failure of the very fast men to avoid accidents is attributed by Wechsler to the fact that very quick men are liable to take chances through overconfidence and hence risk mishap. It is significant that the number of errors, or wrong responses, proved to be a better criterion than simple speed of response in separating the frequent-accident men from those who have few accidents. All men, for instance, who made four or more errors in a fifteen minute test were rated as careless drivers. Frequent-accident men consistently made more errors than few or no-accident men.

The increasing death rate from automobile accidents has directed much study to the problem of locating the accident-prone driver. The method used in picking out the potentially dangerous driver has been to analyze driving skill and construct tests to measure important phases of this complex activity. It is well recognized that a number of factors enter into safe driving. Tests have been developed for measuring the factor of vision—color vision, effects of glare, accuracy in perception of depth, and so on. Also tests are available for measuring hand-eye coördination and other abilities. We have described above one test of reaction time under driving conditions. The attempt—as here—to simulate actual road conditions by means of a miniature situation illustrates a method widely used in vocational psychology. Driving tests are coming more and more into use. When strictly applied such tests have a decided influence in decreasing the accident rate.

During World War II, reaction time tests were included in several batteries designed to select men for special services requiring quick perception and rapid motor response. The Discrimination Reaction Time Test shown in Figure 62 was planned to measure the speed with which an individual can make discrimination reactions to patterns of visual stimuli differing in spatial arrangement. Four stimulus lamps, two red and two green (L_1, L_2, L_3, L_4), set in the form of a square, appeared before the candidate. Candidates reacted by operating one of four toggle switches (S_1, S_2, S_3, S_4) according to definite rules. If lamps L_1 and L_4 are lighted, for instance, the red is "up" with respect to the green and S_1 (the upper switch) must be pushed *up*. If L_3 and L_4 are lighted, the red is to the right of the green and S_2 (switch on right) must be pushed to the *right*.

Other combinations of lamps require the candidate to pull *S3 forward*, and to push *S4 to the left*. If the candidate operates the appropriate switch, the large white signal lamp (which was lighted on each trial) was extinguished immediately, thus telling the candidate that his response was correct. Time taken to operate the correct switch on each series of trials was accumulated on an electric stop clock and determined the candidate's score.

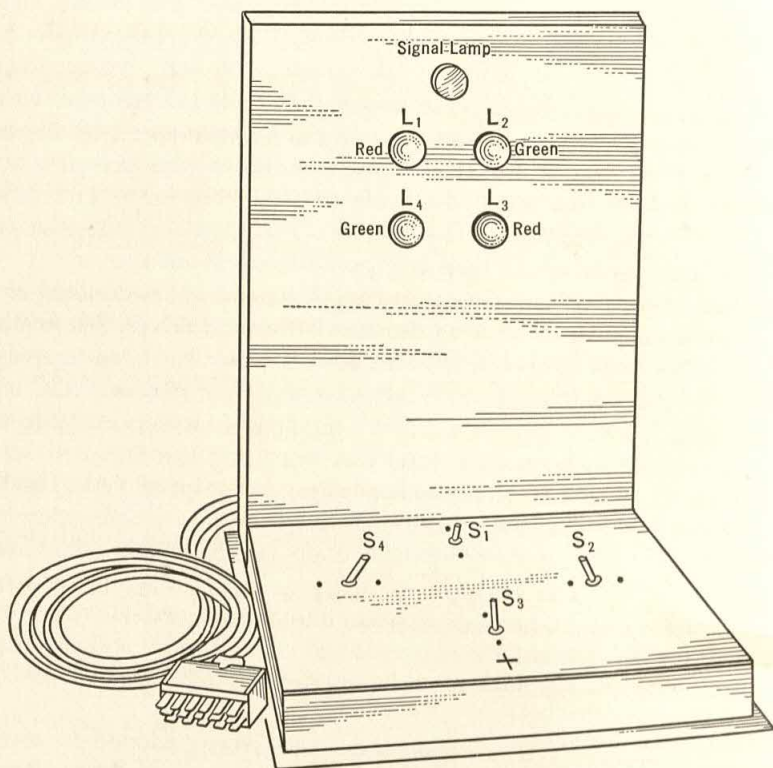


FIG. 62

TEST UNIT OF THE SAM DISCRIMINATION REACTION TIME TEST (CP611D2)
(After Army Air Forces Aviation Psychology Program Research Reports, No. 4, 1947.)

The Discrimination Reaction Time test proved valuable to the Air Forces in the selection of pilots, bombardiers, and navigators as well as in the selection of aerial gunners, radio operator gunners, and flight engineers. It had appreciable correlation (.20 to .40) with success in training and high correlations with paper and pencil tests which were related to training success.

Many of the practical applications of reaction time measurement have occurred since Cattell's time and are the product of new conditions and of an attempt to satisfy new demands. They were stimulated, however, by the early interest of psychologists in general and of Cattell in particular in the objective measurement of the time relations of human performances.

BIBLIOGRAPHY

1. Bingham, W. V., & Freyd, M. *Procedures in employment psychology*. Chicago: A. W. Shaw Company, 1926.
2. Boring, E. G. *A history of experimental psychology*, 2nd ed. New York: Appleton-Century-Crofts, Inc., 1950, chapter 8.
3. Breitwieser, J. V. Attention and movement in reaction time. *Arch. Psychol.*, New York, 1911, 2, No. 18.
4. Cassel, E. E., & Dallenbach, K. M. The effect of auditory distractions upon the sensory reaction. *Amer. J. Psychol.*, 1918, 29, 129-143.
5. Cattell, J. McK. The time taken up by cerebral operations. *Mind*, 1886, 11, 220-242, 377-392, 524-538.
6. Dolley, C. S., & Cattell, J. McK. On reaction-times and the velocity of the nervous impulse. *Nat'l Acad. Science, Memoirs*, 1895, 7, 391-415; also *Psychol. Rev.*, 1894, 1, 159-168.
7. Evans, J. E. The effect of distraction on reaction time. *Arch. Psychol.*, New York, 1916, 5, No. 37.
8. Froeberg, S. The relation between the magnitude of stimulus and the time of reaction. *Arch. Psychol.*, New York, 1907, 1, No. 8.
9. Griffith, C. R. An introduction to applied psychology. New York: The Macmillan Company, 1937, chapter 1.
10. Henmon, V. A. C. The time of perception as a measure of differences in sensations. *Arch. Phil. Psychol. & Sci. Meth.*, 1906, No. 8.
11. Henmon, V. A. C. Professor Cattell's work on reaction time. In the *psychological researches of James McKeen Cattell*. *Arch. Psychol.*, New York, 1914, 4, No. 30, 1-33.
12. Johanson, A. M. The influence of incentive and punishment upon reaction-time. *Arch. Psychol.*, New York, 1922, 8, No. 54.
13. Judd, C. H., McAllister, C. N., & Steele, W. M. Analysis of reaction movements. *Psychol. Monog.*, 1905, 7, No. 29, 141-184.
14. Lauer, A. R. Fact and fancy regarding driver testing procedures. *J. appl. Psychol.*, 1937, 21, 173-184.
15. Lauer, A. R., & Kotvis, H. L. Automotive manipulation in relation to vision. *J. appl. Psychol.*, 1934, 18, 422-431.
16. Lemmon, V. W. The relation of reaction time to measures of intelligence, memory, and learning. *Arch. Psychol.*, New York, 1927, 15, No. 94.
17. Lickley, J. D. *The nervous system*. New York: Longmans, Green & Company, 1931, 15.
18. Melton, A. W. Army air forces aviation psychology program, Research Reports, 1947, No. 4, Apparatus Tests, chapter 6.
19. Mowrer, O. H. Preparatory set (expectancy), Some methods of measurement, *Psychol. Monog.*, 1940, 52, 233.

20. Poffenberger, A. T. Reaction time to retinal stimulation. *Arch. Psychol.*, New York, 1912, 3, No. 23.
21. Shellow, S. M. Selection of motormen: further data on value of tests in Milwaukee. *J. pers. Res.*, 1926, 5, 183-188.
22. Steinman, A. R. Reaction time to change. *Arch. Psychol.*, New York, 1944, 41, No. 292.
23. Toops, H. A., & Haven, S. E. Viewing the traffic problem. *J. appl. Psychol.*, 1937, 21, 185-197.
24. Wechsler, D. Tests for taxicab drivers. *J. pers. Res.*, 1926, 5, 24-30.
25. Woodrow, H. The measurement of attention. *Psychol. Monog.*, 1914, 17, No. 76.
26. Woodrow, H. The faculty of attention. *J. exp. Psychol.*, 1916, 1, 285-318.
27. Woodworth, R. S. *Experimental psychology*. New York: Henry Holt and Company, 1938, chapter 14.

WEBER'S AND FECHNER'S LAWS AND THE DEVELOPMENT OF PSYCHOPHYSICS

1. THE PROBLEMS OF PSYCHOPHYSICS

ALL OF us have noticed that our impressions of differences in the *magnitude*, *extent*, and *amount* of the things we perceive around us do not vary directly with changes in the objects themselves. Every one would agree, for example, that a faint tone does not need to be increased as much as a loud tone for the *change* to be noticed; and that a four-year-old boy does not have to grow as much as his sixteen-year-old brother for the increase in height to be perceived. Furthermore, most of us recognize at once that one pound added to a ten-pound load is more clearly felt as an *increased* weight than one pound added to a fifty-pound load; and that an error of one inch in measuring one foot is far cruder work than an error of one inch in measuring one mile. But few who note these facts have given much thought to the question of how our *perception of a difference*, and the *actual difference itself*, are related, other than to speculate, perhaps somewhat vaguely, that the relationship is clearly not one of simple correspondence.

It so happens that one of the very earliest attempts to employ experimental method in psychology was directed toward solving this problem of the relation of stimulus changes to sensation changes. There were several reasons for this. For one thing the early psychologists were often philosophers, too, and as philosophers were concerned with the problem of how knowledge of the world is obtained through the senses (the old mind-body problem) and with the meaning of experience. Moreover, there already existed a respectable body of fact concerning the sense-organs, the ear, the eye, the skin, which furnished a convenient starting-point for investigation. Whatever the less tangible reasons, the fact remains that the topic of sensation-stimulus dependence bulks large in the history of experimental psychology and has stimulated an enormous amount of research. A whole literature, in fact, has grown up around the subject of *psychophysics*, as this branch of psychology is called; and the purpose

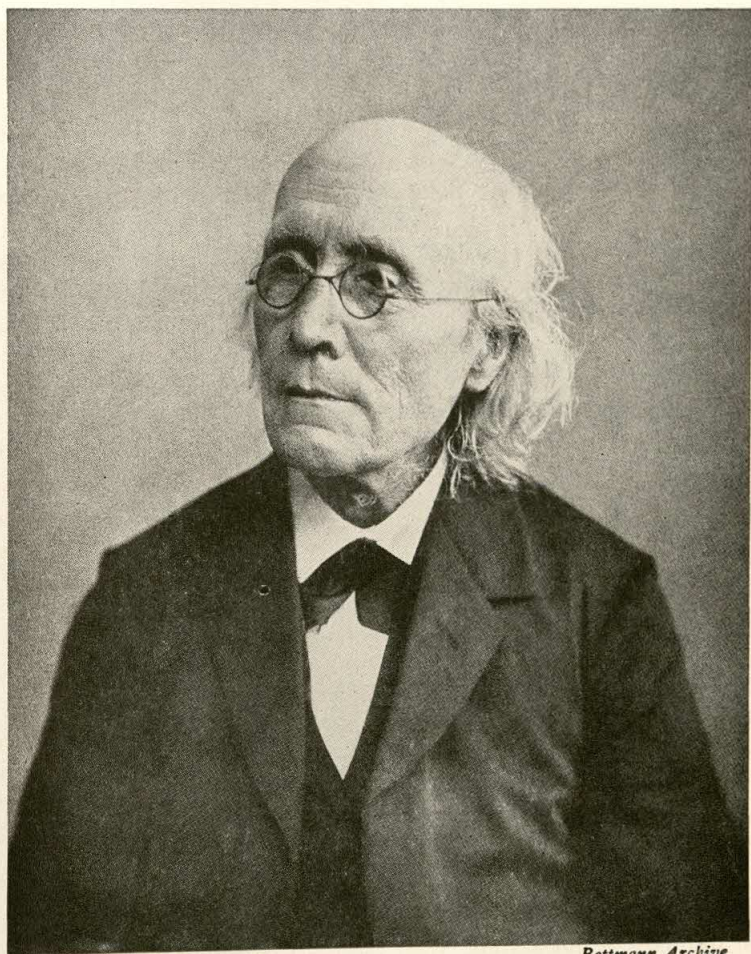
of this chapter is to outline the development of psychophysics and to evaluate in brief space its place in modern psychology.

2. WEBER'S LAW

Psychophysics really began with the work of Ernst Heinrich Weber (1795-1878), although as a separate branch of psychology it did not originate with him. During the years from 1829 to 1834, Weber, who was professor of anatomy in the University of Leipzig, published in Latin a long series of experiments on cutaneous and kinesthetic (or muscular) sensation under the title *De Tactu* (Concerning Touch). Weber was interested in discovering how accurately small differences in weight can be sensed when the weights are lifted by hand or are allowed to rest freely on the surface of the skin. He was interested also in how small a difference between two lines can be ascertained or perceived by the eye. In a series of experiments with weights held between the fingers and lifted by hand, Weber discovered that a weight of 30 ounces could just barely be distinguished as lighter than a weight of 31 ounces, and as heavier than one of 29 ounces. This proportion he found to hold closely for drams¹ as well as for ounces, i.e., 30 drams could just be felt as lighter than 31 and heavier than 29 drams. With practiced subjects, the same proportion held for lighter weights—14.5 ounces or 14.5 drams, for instance, could just be distinguished from 15 ounces or 15 drams, respectively. Smaller differences than these, said Weber, are very rarely "sensed," while greater differences are readily discriminated. In experiments concerned with judging the length of lines, Weber discovered that a different but consistent principle of proportion held good. A line of 101 millimeters could just be discriminated as longer than one of 100 millimeters, while a line of 51 millimeters was judged to be just observably longer than one of 50.5 millimeters. A stimulus difference which can just be perceived is called a "difference limen" or DL.

On the basis of his results, Weber formulated the famous generalization known subsequently as Weber's law. In comparing objects, he says, it is clear that we perceive not the *actual difference* between the two objects, but the *ratio* of this difference to the magnitudes of the two objects compared. To put it in other words, the observed differences between two objects is not absolute and completely independent of the objects themselves, but is relative to the *size* of the stimuli and is a *constant fraction* of one of them, the so-called "standard stimulus." This constant fraction must be discovered by experiment. For weights lifted by hand, Weber found it to be $\frac{1}{30}$ to $\frac{1}{40}$, and for lines approximately

¹ A dram = $\frac{1}{16}$ ounce avoirdupois.



Bettmann Archive

GUSTAV THEODOR FECHNER

1801-1887

in 75 per cent (or other designated per cent) of the trials is a constant fraction of the size of the stimulus.

3. FECHNER'S CONTRIBUTIONS TO PSYCHOPHYSICS: THE WEBER-FECHNER LAW

The equation $\frac{dR}{R} = C$ was Weber's final statement of his law, and constitutes his chief contribution to the study of the relation between sensory judgments and stimulus intensities. It remained for Gustav Theodor Fechner (1801-1887) to take up where Weber left off and, building upon Weber's law, erect the intricate structure called psychophysics. Fechner, besides being professor of physics at Leipzig was also a philosopher and something of a mystic. In Weber's generalization he saw a means of studying quantitatively the relation between the physical and mental worlds. In his *Elements of Psychophysics*, first published in 1860, Fechner defined psychophysics as "an exact science of the functional relation or relations of dependency between body and mind." The physical world for Fechner was represented by the physical stimuli, and the psychical world by the sensations within the organism aroused by these stimuli.

Fechner laid down two principles which enabled him to amplify and extend Weber's law. First, he assumed that a large sensation may be thought of as the *sum* of a number of small sensations; and secondly he assumed that just noticeable differences (*j.n.d.'s*) in sensation are equal, and hence are suitable units for measuring sensation changes. Accepting Weber's equation $\frac{dR}{R} = C$ as fundamentally correct, Fechner on the basis of experiments extending over many years,³ reformulated Weber's law to read as follows: *When stimuli increase by a constant ratio, the sensations aroused by them increase by equal increments or steps.* What Fechner meant by this statement, and how it depends upon his assumptions may be seen more clearly, perhaps, in Figure 63.

In the diagram we begin with *two* stimuli, a standard and a variable comparison-stimulus, each designated by the value 100 on the horizontal scale. These two stimuli may be thought of as two lights of brightness value 100, or two lines each 100 mms. long, or two tones each of an absolute intensity designated by 100. To represent these values on the stimulus scale (the horizontal line), we lay off an arbitrary distance from 0 to the point marked 100. Then on the vertical line above 100 (R_1) we lay off an arbitrary distance to represent the strength of the sensation (S_1)

³ During the years from 1855 to 1859, Fechner alone made 67,072 weight comparisons, acting both as experimenter and as subject.

produced by stimulus 100. Suppose now that we increase very slowly the comparison stimulus 100, and that when this stimulus reaches the value 110 (R_2) it is judged to be *just noticeably greater* than the standard stimulus, which remains, of course, at 100 or R_1 . The sensation value S_2 (on the vertical line) corresponding to stimulus 110 (R_2) will now equal the sensation S_1 (corresponding to stimulus 100) plus one *j.n.d.* unit—on Fechner's assumption that the *j.n.d.* may be taken as the unit of sensation-change. In Figure 63 the *j.n.d.* is taken as a small and arbitrary increment which when added to S_1 gives S_2 .

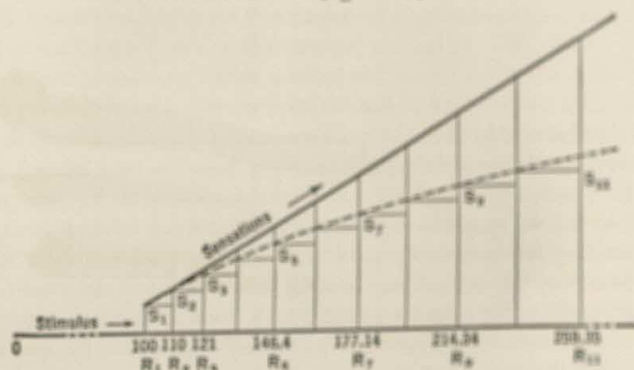


FIG. 63

GRAPHIC REPRESENTATION OF THE RELATION BETWEEN INCREASES IN THE INTENSITY OF THE STIMULUS AND INCREASES IN THE INTENSITY OF THE SENSATION

Now let us begin with our two stimuli, the standard and comparison, both at the value 110 on the scale, and increase the comparison stimulus slowly until it becomes again just noticeably greater than the constant standard, still at 110. This point should be reached theoretically at 121 ($110 + \frac{110}{10}$), since in the previous trial our Weber fraction was found to be $\frac{1}{10}$. The sensation S_2 corresponding to the stimulus 121 (R_2) is S_1 plus one *j.n.d.*, or S_1 plus two *j.n.d.*'s and accordingly the distance S_2 plus two *j.n.d.* units is laid off above stimulus 121 (R_2). The other stimuli in Figure 63, viz., 133.1 (R_3), 146.4 (R_4), 161.04 (R_5), have all been calculated in the same manner as 121 and 110—in each case *one tenth* of the preceding stimulus magnitude has been added on to the stimulus scale to give a just noticeable increase on the sensation scale and to accord with Weber's law. The sensation (vertical distance) corresponding to the stimulus on the baseline is in every instance equal to the sensation aroused by the preceding stimulus plus one *j.n.d.* unit.

It should be noted that the stimuli in Figure 63 increase in *geometrical progression* as we go from 100 to 161.04, while the sensations, the psychical

in 75 per cent (or other designated per cent) of the trials is a constant fraction of the size of the stimulus.

3. FECHNER'S CONTRIBUTIONS TO PSYCHOPHYSICS: THE WEBER-FECHNER LAW

The equation $\frac{dR}{R} = C$ was Weber's final statement of his law, and constitutes his chief contribution to the study of the relation between sensory judgments and stimulus intensities. It remained for Gustav Theodor Fechner (1801-1887) to take up where Weber left off and, building upon Weber's law, erect the intricate structure called psychophysics. Fechner, besides being professor of physics at Leipzig was also a philosopher and something of a mystic. In Weber's generalization he saw a means of studying quantitatively the relation between the physical and mental worlds. In his *Elements of Psychophysics*, first published in 1860, Fechner defined psychophysics as "an exact science of the functional relation or relations of dependency between body and mind." The physical world for Fechner was represented by the physical stimuli, and the psychical world by the sensations within the organism aroused by these stimuli.

Fechner laid down two principles which enabled him to amplify and extend Weber's law. First, he assumed that a large sensation may be thought of as the *sum* of a number of small sensations; and secondly he assumed that just noticeable differences (*j.n.d.'s*) in sensation are equal, and hence are suitable units for measuring sensation changes. Accepting Weber's equation $\frac{dR}{R} = C$ as fundamentally correct, Fechner on the basis of experiments extending over many years,³ reformulated Weber's law to read as follows: *When stimuli increase by a constant ratio, the sensations aroused by them increase by equal increments or steps.* What Fechner meant by this statement, and how it depends upon his assumptions may be seen more clearly, perhaps, in Figure 63.

In the diagram we begin with *two* stimuli, a standard and a variable comparison-stimulus, each designated by the value 100 on the horizontal scale. These two stimuli may be thought of as two lights of brightness value 100, or two lines each 100 mms. long, or two tones each of an absolute intensity designated by 100. To represent these values on the stimulus scale (the horizontal line), we lay off an arbitrary distance from 0 to the point marked 100. Then on the vertical line above 100 (R_1) we lay off an arbitrary distance to represent the strength of the sensation (S_1)

³ During the years from 1855 to 1859, Fechner alone made 67,072 weight comparisons, acting both as experimenter and as subject.

produced by stimulus 100. Suppose now that we increase very slowly the comparison stimulus 100, and that when this stimulus reaches the value 110 (R_2) it is judged to be *just noticeably greater* than the standard stimulus, which remains, of course, at 100 or R_1 . The sensation value S_2 (on the vertical line) corresponding to stimulus 110 (R_2) will now equal the sensation S_1 (corresponding to stimulus 100) *plus one j.n.d. unit*—on Fechner's assumption that the *j.n.d.* may be taken as the unit of sensation-change. In Figure 63 the *j.n.d.* is taken as a small and arbitrary increment which when added to S_1 gives S_2 .

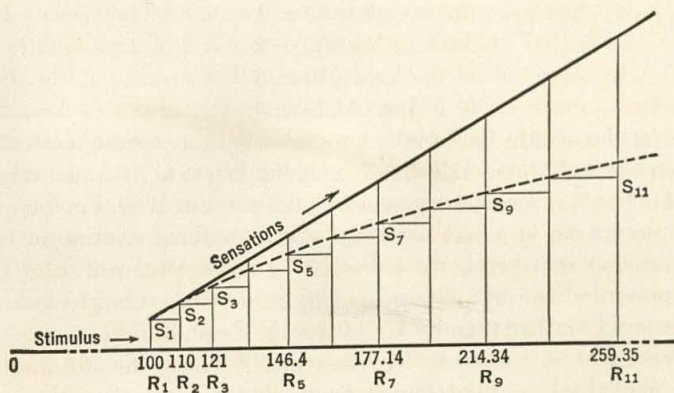


FIG. 63

GRAPHIC REPRESENTATION OF THE RELATION BETWEEN INCREASES IN THE INTENSITY OF THE STIMULUS AND INCREASES IN THE INTENSITY OF THE SENSATION

Now let us begin with our two stimuli, the standard and comparison, both at the value 110 on the scale, and increase the comparison stimulus slowly until it becomes again just noticeably greater than the constant standard, still at 110. This point should be reached theoretically at 121 ($110 + \frac{11}{10}$), since in the previous trial our Weber fraction was found to be $\frac{1}{10}$. The sensation S_3 corresponding to the stimulus 121 (R_3) is S_2 plus one *j.n.d.*, or S_1 plus two *j.n.d.*'s and accordingly the distance S_1 plus two *j.n.d.* units is laid off above stimulus 121 (R_3). The other stimuli in Figure 63, viz., 133.1 (R_4), 146.4 (R_5), 161.04 (R_6), have all been calculated in the same manner as 121 and 110—in each case *one tenth* of the preceding stimulus magnitude has been added on to the stimulus scale to give a just noticeable increase on the sensation scale and to accord with Weber's law. The sensation (vertical distance) corresponding to the stimulus on the baseline is in every instance equal to the sensation aroused by the preceding stimulus plus one *j.n.d.* unit.

It should be noted that the stimuli in Figure 63 increase in *geometrical progression* as we go from 100 to 161.04, while the sensations, the psychical

series, increase in *arithmetic progression*.⁴ Whenever we have two series which correspond point by point, the one increasing geometrically, the other arithmetically, we know from mathematics that the relation between the two variables must be logarithmic. Hence the sensation values (AP) may be here thought of as the logarithms of their corresponding stimulus-values (GP). The logarithmic relationship between a geometric and an arithmetic series may be simply illustrated by the common logarithms of 10 and its multiples. From any table of logarithms we find that

$$\log 1 = 0$$

$$\log 10 = 1$$

$$\log 100 = 2$$

$$\log 1,000 = 3$$

$$\log 10,000 = 4$$

On the *left-hand* side the numbers increase by a constant multiple (10), while on the *right-hand* side the logarithms increase by a constant increment (1). The first series is a geometric, the second an arithmetic, progression. If we let our numbers represent stimuli and our logs represent their corresponding sensations, we have the situation pictured in Figure 63. If the sensation values had increased in *direct proportion* to the increase in the stimuli (rather than by equal steps), we should have obtained the *linear* relationship represented by the rapidly rising straight line instead of the more slowly rising logarithmic curve shown in the diagram.

Discovery of the logarithmic relationship between stimuli and their sensations led Fechner to restate Weber's law in mathematical terms as follows: *Sensations are proportional to the logarithms of their exciting stimuli, or, in the form of an equation:*

$$S = C \log R$$

wherein S is the sensation, R the stimulus, and C a constant to be determined from the experiment. This equation is generally known as the Weber-Fechner law or simply as Fechner's law. The first two quantities, of course, are variables, the third, C , is fixed for a given series of stimuli, e.g., weights, lines, or brightnesses, but varies from one sense modality to another. C 's value depends partly upon the sense modality, the precision of measurement, and partly upon the choice of the zero point for sensation intensities.

⁴ A geometric progression (GP) is a series in which each term (except the first) is derived from the preceding by multiplying by a constant number or *ratio*: e.g., 4, 12 (4×3), 36 (12×3), 108 (36×3). In the present case we have 100, 110 ($100 \times 1\frac{1}{10}$), 121 ($110 \times 1\frac{1}{10}$), 133.1 ($121 \times 1\frac{1}{10}$), 146.4 ($133.1 \times 1\frac{1}{10}$). An arithmetic progression (AP) is a series in which each term (except the first) is derived from the preceding by the *addition* of a constant number: e.g., 3, 7 ($3 + 4$), 11 ($7 + 4$), 15 ($11 + 4$), 19 ($15 + 4$). In Figure 63 we have S_1 , S_2 ($S_1 + 1$ j.n.d.), S_3 ($S_2 + 1$ j.n.d.), S_4 ($S_3 + 1$ j.n.d.), S_5 ($S_4 + 1$ j.n.d.), S_6 ($S_5 + 1$ j.n.d.).

Fechner's logarithmic law is regarded by many psychologists as one of the most important generalizations in psychology. This principle, to be sure, is an extension and modification of Weber's law, but it is also much more than this. Weber's equation, $\frac{dR}{R} = C$, said nothing about sensation intensities; it stated simply that, for a given proportion of correct judgments, the ratio of the stimuli is constant, no matter what the absolute magnitudes of the objects compared. Fechner's equation, on the other hand, expresses a *functional* relationship between physical stimuli and their corresponding sensation intensities—a relationship which Fechner, at least, believed to hold true over the whole range of perceptible stimuli.

4. LIMITATIONS OF WEBER'S AND FECHNER'S LAWS

As an expression of the general law of the relativity of all judgments, or the dependence of our perceptions of change upon the magnitude of the thing changed, Weber's law serves to describe quite well many facts of everyday perception. Familiar examples (see also p. 331) readily come to mind: An inch is more perceptible when added to a man's finger than when added to his height; a room lighted by an electric light is scarcely brightened at all by the addition of a candle, but when lighted by only a single candle is markedly brightened by the addition of another candle; to hear a pin drop the room must be very quiet; five pounds added to a baby's weight represents a much more noticeable change than five pounds added to a man's weight. These illustrations show the soundness of Weber's principle, but there are distinct limitations to the law which must be recognized in trying to evaluate it. In the first place, Weber's law holds only approximately within *any* field. This is true whether we are concerned with the perception of linear magnitudes (e.g., lengths of lines), the perception of weights by passive pressure on the skin or by active lifting, the intensity of lights and sounds, the judgment of duration and movement, or the perception of pitch differences. Secondly, Weber's law holds only in the *middle range* of stimulus intensities within a given sense modality. It breaks down with very weak or very strong stimuli. Expressed differently, the ratio $\frac{dR}{R}$ remains fairly constant over the middle range of stimuli, but increases markedly when the stimuli become very strong or very weak.

A series of careful experiments which show this phenomenon clearly were performed by König and Brodhun in 1888–1889. These investigators found the DL for light intensities varying from very faint to very bright

illumination to be a constant fraction (approximately .020) of the stimulus over a wide range of medium brightnesses. The Weber fractions were considerably larger, however, for very faint and for very bright lights than for medium intensities. This may be seen in König and Brodhun's data, some of which are reproduced in Table XXXV below.

TABLE XXXV

The stimuli are given in an arbitrary unit. The *ratios* give the fraction which the just-noticeable difference is of the absolute stimulus. Note that from about 400 to 100,000 the ratios are approximately equal. (After König and Brodhun)

Stimulus:	1	2	4	10	20	40	100	200	400	1,000
Ratio:	.256	.175	.120	.070	.048	.037	.030	.025	.022	.020
Stimulus:	2,000	4,000	10,000	20,000	40,000	100,000	200,000	400,000		
Ratio:	.018	.017	.016	.017	.018	.021	.029	.038		

These variations in the Weber ratio are attributable in some degree to adaptation and various slight disturbances. But such factors are not sufficient, in a careful experiment, to account for the wide discrepancies throughout the range of intensities. We are forced to conclude therefore that Weber's law breaks down at the extremes of the stimulus scale.

Does Fechner's law fare any better than Weber's as a universal principle? Again the answer is not a simple yes or no. In the first place, Fechner's law, like Weber's, holds only approximately; and like Weber's holds only over the middle range of stimuli. In the second place, Fechner's two assumptions have not stood the test of later experiment, nor of adverse criticism. Psychologists have not been slow to point out that a large sensation is not psychologically the sum of many elementary sensations, but is a new experience, just as water is not simply the sum of oxygen and hydrogen, but is essentially a new product. Moreover, the *j.n.d.* is by no means a fixed unit. The just noticeable difference between two lights or two tone intensities will vary from one observer to another, and for the same observer from time to time. Theoretically, there is no such thing as a "just noticeable difference," since even very small differences will occasionally be sensed, while large differences will sometimes not be perceived. A third objection to the universality of Fechner's law is based upon the difficulty of fixing definitely the point of zero sensation. We know from mathematics that the logarithm of 1 is zero. Fechner fitted this into his formula, $S = C \log R$, by assigning the value 1 to that stimulus which arouses a just barely perceptible (zero) sensation. But the eye, the ear, and the skin rapidly become adapted to faint lights, faint sounds, and faint touches; and this adaptation, plus fatigue, slight distractions, and lapses of attention, shift and distort the zero point of sensation. Thus it

would seem that the variability and instability inherent in the sensory receptors themselves must limit the universality of Fechner's law.

These objections to Fechner's law limit the range of its applicability. Nevertheless the principle of logarithmic relationship between stimulus changes and our perceptions of these changes has been found useful in many problems of biology, physics, and engineering. Several illustrations may be given. The relation of visual acuity to the brightness of the field is logarithmic. If absolute acuity (A) "is the reciprocal of the smallest visual angle for which neighboring contrasted portions of the field can be separated," and if B is the brightness of the field, then $A = c + k \log B$ (ref. 13). The constants c and k are dependent upon the units employed, the character of the field, and the eye itself. This equation is Fechner's law, $S = C \log R$; it holds over the middle range of intensities. In the field of biophysics, it has been shown in certain organisms that the relationship between the reciprocal of the *latent time* of response (e.g., reaction time to light) and the *intensity* of stimulation is logarithmic, for a constant exposure time. If we let $X = 1/\text{latent time}$, and $I = \text{intensity}$, then—over the middle range of intensities— $X = k \log I$, which again is Fechner's law.

Telephone engineers have made good use of Fechner's principle. Experiments in audition have shown that the relation between "loudness" as sensed by the ear and the physical intensity (vibration amplitude) of the sounding body is logarithmic. This relation can be put in the form of an equation, $a = \log J$, wherein $a = \text{loudness measured in "bels"}$ and J is the physical intensity of the sound in microwatts. A bel (so-called after Alexander Graham Bell) equals 10 decibels, the decibel being the smallest change in loudness which the ear can detect. A decibel or "sensation unit" is thus clearly analogous to the Fechnerian *j.n.d.* or to the DL for intensity of sound. The decibel is the unit most often used by audition experts in this country.

From these illustrations it would appear that Fechner's law is by no means a dead issue, although it is a limited principle rather than the universal generalization Fechner proclaimed it to be. As far as psychology is concerned, the present worth of Fechner's law would seem to be that it serves to represent quantitatively the stimulus-response relation over the middle range of stimulus intensities.

5. INTERPRETATIONS OF THE WEBER AND FECHNER LAWS

The interpretation of Fechner's law has led to much discussion and many theories. Fechner, as we have seen, regarded his law, $S = C \log R$, as essentially an equation expressing the fundamental relationship between

In determining the just-noticeable difference in brightness or the threshold of sensitivity, there are four stages. (1) We begin with the two windows *P* and *Q* of equal brightness, and slowly increase the brightness of *Q* until it is judged to be just perceptibly brighter than *P*. Call this point *A*. (2) Next, with *Q* much brighter than *P*, we decrease the brightness of *Q* (the variable) until it is judged to be of *equal* brightness with (or not perceptibly different from) *P*. Call this second determination *B*. (3) As in (1) above, we begin again with the windows *P* and *Q* of the same brightness, but this time we decrease *Q*'s brightness very slowly until the observer judges it to be just less bright (darker) than *P*. This is point *C*. (4) Finally, following (2) above, we begin with *Q* much darker than *P* and increase its brightness until the two stimuli are judged to be equal. This is point *D*. From the average of these four points, *A*, *B*, *C*, and *D*, we can obtain the "general threshold," i.e., the brightness-difference which is just capable of being perceived. Needless to say, observations must be repeated again and again to insure reliable results.

Table XXXVI illustrates the method of minimal change in an experiment similar to the one outlined above. Note that the ten trials taken under the four different sets of conditions have been averaged separately, and that they show considerable variation the one from the other. If we average the four determinations, *A*, *B*, *C*, and *D*, the result, 4.2, gives the general threshold, the average just observable difference, or the average DL. This value is a measure of the observer's keenness of discrimination under the given conditions. Other interesting comparisons may be made. By averaging *A* and *C*, for instance, and then *B* and *D*, we can compare the accuracy of perception when the change is from equality to difference with that when the change is from difference toward equality. When we do this, it is evident that judgment was more accurate when the difference between the standard and the comparison was *decreasing* than when it was *increasing*. The upper threshold, i.e., $\frac{A+B}{2}$, is determined from those settings in which the comparison stimulus is *brighter* than the standard; and the lower threshold, i.e., $\frac{C+D}{2}$, is determined from those settings in which the comparison is *darker* than the standard. It is clear from Table XXXVI that the upper threshold was considerably smaller than the lower threshold, and hence perception of brightness changes was more accurate under the first set of conditions than under the second.

2. *The Method of Constant Stimuli.* Like the method of minimal change, this method was also developed from an earlier procedure, the *method of right and wrong cases*. Though it may be employed with other stimuli, as for example, linear extents (lines), this method has most often been

TABLE XXXVI

DETERMINATION OF THE JUST NOTICEABLE DIFFERENCE IN LIGHT INTENSITY
BY THE METHOD OF MINIMAL CHANGE

The standard was a four candle-power lamp, and ten trials were taken at each of four stages. In one half of the trials at each stage the standard was on the right, and in one half, on the left. (Data adapted with some changes from the original by the writer)

	<i>Average Setting</i>
(A) From equality upward, i.e., the comparison stimulus becomes brighter	4.6
(B) Toward equality downward, i.e., the comparison stimulus becomes darker	2.7
(C) From equality downward, i.e., comparison becomes darker	5.6
(D) Toward equality upward, i.e., comparison becomes brighter	3.9
General threshold $\frac{(A + B + C + D)}{4}$	4.2
Average increasing difference $\frac{(A + C)}{2}$	5.1
Average decreasing difference $\frac{(B + D)}{2}$	3.3
Upper threshold $\frac{(A + B)}{2}$	3.7
Lower threshold $\frac{(C + D)}{2}$	4.8

used to measure tactual and kinesthetic sensitivity through the lifting of weights, and hence we shall illustrate it with lifted weights data. (See Table XXXVII.) Suppose that we have selected a standard weight of 100 grams, and comparison weights of 88, 92, 96, 104, 108, and 112 grams. The problem is to determine the amount by which the 100-gram standard must be increased or decreased for the change to be just observable. But instead of changing the standard by small amounts we compare it over and over again, first with the lighter and then with the heavier weights, noting in each case the number of correct judgments. Judgments may be recorded as heavier, as lighter, or as equal. To illustrate, in comparing 100 grams and 108 grams, the judgment "heavier" might be given 80 per cent of the time, "equal" 12 per cent, and "lighter" 8 per cent of the time; and such results as these will be secured with each weight pair, 100-88, 100-104, 100-112, and so on. When there are an equal number of heavier and lighter comparison stimuli, the standard weight is usually lifted first and the comparison weight second, the judgment of heavier, lighter, or equal expressing the subject's opinion of the relation between

the variable (the second) weight and the standard (the first) weight. That weight above 100 grams which gives just 50 per cent "heavier" judgments determines the *upper threshold*, or the just-perceptibly-heavier difference; while that weight below 100 grams which gives just 50 per cent "lighter" judgments determines the *lower threshold*, or the just-perceptibly-lighter difference. The average of these two values is the *general threshold*, or the average difference limen. It is seldom that a weight difference will be found experimentally which gives exactly 50 per cent "heavier" or 50 per cent "lighter" judgments, so that interpolation is nearly always necessary. The two 50 per cent points may be determined fairly accurately from graphs, and somewhat more precisely from formulas for interpolation between the percentages actually found.

The procedure for the method of constant stimuli above outlined is the older, or classical, form in which the method was developed by Fechner, Müller, and Urban. It is an accurate and precise method when used with highly trained subjects. But with those less well trained it is open to considerable objection, because of the allowance of equal judgments. We have no way of controlling the number of judgments which a subject will call "equal," and hence if he is extremely cautious, a large part of his data will be so classified, and hence not directly usable.

In 1892 Fullerton and Cattell showed in a series of experiments that if a subject will *guess* when inclined to say equal, he is more often right than wrong, as there is usually present some slight basis for a judgment, although it may be too small to be very convincing. These investigators simplified the method of constant stimuli, permitting only *two* judgments—"heavier" or "lighter." They made it up, in a sense, to the subject for thus forcing his answer by allowing him to qualify each judgment by expressing his degree of confidence as "sure," "fairly sure," "a guess."

In more recent experiments, Kellogg found that the *variability* of his subjects' responses differed little whether a two or three category judgment was allowed. Hence the two-answer scheme seems justified by experimental results. In this simpler form of the method of constant stimuli, that weight difference which gives 75 per cent "heavier" judgments determines the upper threshold, while that difference giving 75 per cent "lighter" judgments fixes the lower threshold. The 75 per cent threshold is arbitrary, to be sure (p. 346), but it is, perhaps, the most reasonable value. The average of the upper and lower thresholds gives the general threshold, or average DL. This DL, it will be noted, cannot properly be called a *j.n.d.* unless we assume that the difference which is perceived correctly in 75 per cent of the trials is in fact "just noticeable." As this is often questionable it is more accurate to call such a DL the "75 per cent threshold."

Certain data selected from an experiment on weight-lifting are given in Table XXXVII to illustrate the method of constant stimuli in its simpler form. In all, three hundred comparisons were made for each weight-pair, e.g., 100-96, 100-100, 100-112, so that the percentages of correct ("heavier" and "lighter") judgments were quite reliably determined. The standard weight throughout the experiment was 100 gms; and the comparison 100-100 was included as a check. It will be seen in Table XXXVII that the lower 75 per cent threshold is exactly 8 gms, since the comparison 100-92 gave just 75 per cent correct judgments. Unfortunately, the upper 75 per cent threshold cannot be so readily obtained. It is clear, however, that it must lie between 4 and 8 gms, since the weight-pair 100-104 gave 66 per cent correct judgments and the pair 100-108, 85.3 per cent correct judgments. By simple interpolation, the value most likely to give 75 per cent correct falls almost exactly at 106, just half-way between 104 and 108.

TABLE XXXVII

DISCRIMINATION OF LIFTED WEIGHTS: METHOD OF CONSTANT STIMULI

The combined records for six observers, each of whom made 50 comparisons of each weight pair, given 300 comparisons for each pair, and 1,800 for all six pairs.

Comparison Weights	Percentage of Judgments "Heavier"	Percentage of Judgments "Lighter"
88	15.7	84.3
92	25.0	75.0
96	43.7	56.3
100	53.3	46.7
104	66.0	34.0
108	85.3	14.7
112	90.7	9.3
Upper 75% threshold	6 gms.	
Lower 75% threshold	8 gms.	
General 75% threshold	7 gms.	

The results of Table XXXVII are presented graphically in Figure 64. The two curves showing the frequency of heavier and lighter judgments for each comparison weight (including 100 gms) are called *psychometric functions*. The points where these two curves cut the 75 per cent horizontal determine the upper and lower thresholds. The first is at 106 and the second at 92. An interesting value is the point of *subjective equality*. This is found where the two psychometric curves cross the 50 per cent horizontal: it tells us what comparison weight should yield just 50 per cent heavier and 50 per cent lighter judgments. It is evident from our present data that the point of subjective equality is at 98.5 gms, approximately, instead of 100 gms—that half of the time 98.5 gms would be judged

heavier and half of the time lighter than 100 gms. This indicates a constant tendency for our observers to *overestimate* the comparison weights, a tendency shown also by the greater proportion of heavier judgments on 88, 92, and 96 than lighter judgments on 112, 108 and 104.

If we put the lower threshold at 8 grams and the upper threshold at 6 grams, the general threshold is the arithmetic mean of these two values, or 7 grams. This means that, in many trials, the difference between 100 grams and 107, or 100 grams and 93, should be correctly perceived 75 per cent of the time. In order to verify Weber's law for these data, it would be necessary to show that for a standard of 200 grams, the 75 per cent threshold is *fourteen* grams; for a standard of 500 grams, *thirty-five* grams, and so on. That is, the DL in each case must be a constant fraction of the standard stimulus.

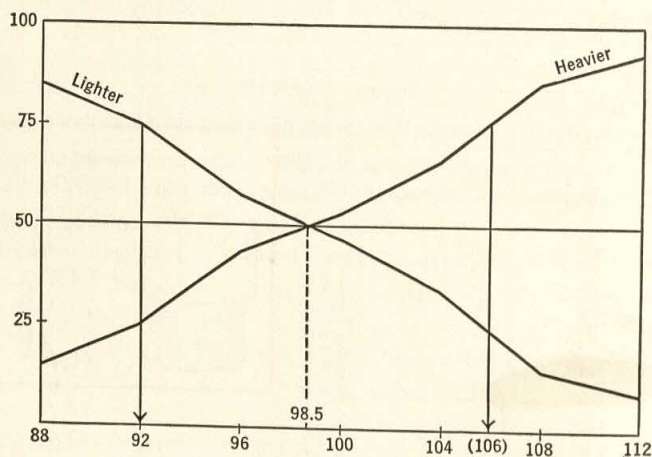


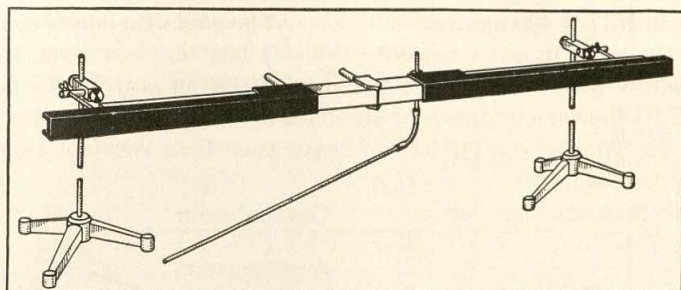
FIG. 64

PSYCHOMETRIC FUNCTIONS OBTAINED WITH TWO JUDGMENT CATEGORIES

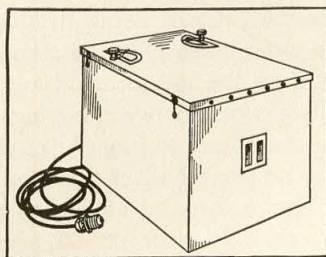
3. *The Method of Average Error.* This method is based on the assumption that measurements of just noticeable differences are essentially measurements of observational errors, i.e., of the limitations or lack of sensitivity of the subject's sensory and nervous mechanism. Ordinarily a subject is given a fixed standard stimulus and an adjustable or variable one which he is instructed to make equal to the standard. The amount by which the observer misses this standard is a measure of his "error of observation."

We may illustrate this method with the Galton bar, often used in the measurement of just non-perceptible differences between linear magnitudes. The Galton bar consists of a strip of flat enameled wood about 2.5 centimeters in height and 100 centimeters in length. It is divided in

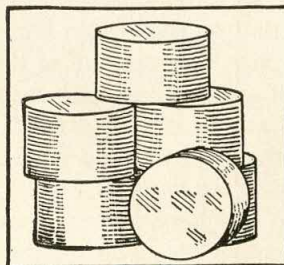
the center by a small wedge, and on the reverse side contains a graduated scale, by means of which the length of each half can be measured. Two sliding sleeves of metal permit varying distances to be set off from the center or 0-point. The whole apparatus is mounted against a black background. (See Fig. 65.) Now suppose a standard length of twenty centimeters is set off on the *right-hand* half of the bar and the subject is in-



A



B



C

FIG. 65

SOME APPARATUS USED IN PSYCHOPHYSICAL EXPERIMENTS

- A. Galton bar used in studying the accuracy of estimates of visually perceived lengths.
 B. A photometer often used in measuring brightness discrimination.
 C. Weights used in experiments on discrimination, or in studies of the kinesthetic sense.

structed to adjust the slide on the *left-hand* half from a much smaller setting than twenty centimeters up to apparent equality with this standard. Twenty trials at least should be made, and the average amount of error (overestimation or underestimation) calculated. A second series of trials is then taken from a too-large setting (greater than twenty centimeters) down to equality with the standard, and the average error again found. The whole procedure must be repeated with the standard set off on the left and the variable on the right half of the bar. The amount of variation or fluctuation in the subject's settings, the amount by which he misses the standard, gives a measure of his *average error*. The tendency of the sub-

ject repeatedly to overestimate or underestimate the standard is called his *constant error*.

TABLE XXXVIII

DATA FOR EIGHTY TRIALS ON THE GALTON BAR, FORTY WITH STANDARD ON THE LEFT AND FORTY WITH STANDARD ON THE RIGHT

(Adapted by writer with several small changes)

<i>Standard left—forty trials</i>		<i>Standard right—forty trials</i>	
	<i>cms.</i>		<i>cms</i>
From too small	18.5	From too small	18.7
From too large	19.2	From too large	19.8
General average	19.05		
Average (from too-small settings)	18.6		
Average (from too-large settings)	19.5	Constant error	.95
		Average error	1.04

Table XXXVIII illustrates the method of average error with data for eighty trials on the Galton bar, the standard length being twenty centimeters. Note that one half of the trials were taken with the standard on the right and one half with the standard on the left to avoid any "place error." Each set of forty trials is equally divided between settings which are, on the one hand, too small and on the other too large; viz., one half of the time the variable was set shorter and one half of the time longer than the twenty-centimeter standard. The average of all eighty trials is 19.05 centimeters, which indicates a *constant error* of .95 centimeter (20—19.05). This is the amount by which, on the average, the observer fell short of the standard, or the amount by which he *overestimated* his settings. The extent to which the observer, on the average, missed the standard is given by his *average error* of 1.04 centimeters. To calculate the AE (average error), we sum up all of the deviations made in the eighty trials from the twenty-centimeter standard and divide by the number of trials. The calculation of the AE is not shown in the table, but can easily be made from the original data. To verify Weber's law, the AE found for a forty-centimeter standard should be twice as large as the AE for our twenty-centimeter standard.

When the observer made his settings from the too-small position (average 18.6), he fell short of the standard to a greater degree than when he made his adjustments from a too-large position (average 19.5). Obviously in both cases he overestimated his settings, however, since both averages are less than the twenty-centimeter standard.

7. THE PRESENT STATUS OF PSYCHOPHYSICS

The evaluation of Fechner's work is not an easy task. Many and various opinions have been expressed regarding it, from the enthusiastic commendation of Titchener, who regarded Fechner as the father of mental measurement, to the caustic remarks of William James, who wrote: "Fechner's book was the starting point of a new department of literature, which it would be perhaps impossible to match for qualities of thoroughness and subtlety, but of which, in the humble opinion of the present writer, the proper psychological outcome is just *nothing*." As is usually the case, the truth is somewhere between these two rather extreme views, but the judgment of time inclines toward Titchener's rather than James's view.

It must be admitted at the outset that much of Fechner's psychophysics has not stood the test of research. His concept of a large sensation as the sum of many smaller sensations, as well as his idea that *j.n.d.*'s are equal over the scale of intensities, have been rejected by present-day psychologists. Also his psychophysical law has been found to have limited instead of, as he thought, universal application. Modern psychology is interested more in the larger aspects of behavior than in the study of sensation and small sensory changes. The development of mental tests with their emphasis upon individual differences has, no doubt, had much to do with this shift in orientation, while the increasing emphasis upon the *wholeness* of the organism's activities, has played a not unimportant rôle. The psychologist of today is interested in the psychophysical methods as tools to be used in the study of problems involving the determination of sensory acuity, sensory efficiency, and the accuracy and precision of movement. Or he may be interested in the possibility of adapting these techniques more directly to problems of mental measurement and practical use. Rarely is he interested in Weber's or Fechner's principles as such. To illustrate, in 1910 E. L. Thorndike devised his Handwriting Scale, the first scale for the measurement of an educational product, proceeding on the assumption that equally often noticed differences between specimens of handwriting are equal. According to this principle, if the difference in excellence between handwriting specimens A and B is noted by competent judges as often as is the difference between specimens X and Y, then the difference in quality between A and B is equal to the difference between X and Y. Thorndike's assumption, as well as his technique, are direct outgrowths from the psychophysical methods. More recently, Thurstone has adapted and extended the psychophysical concepts on the theoretical side and has applied them to the problem of scale-making and to the study of attitudes and preferences.

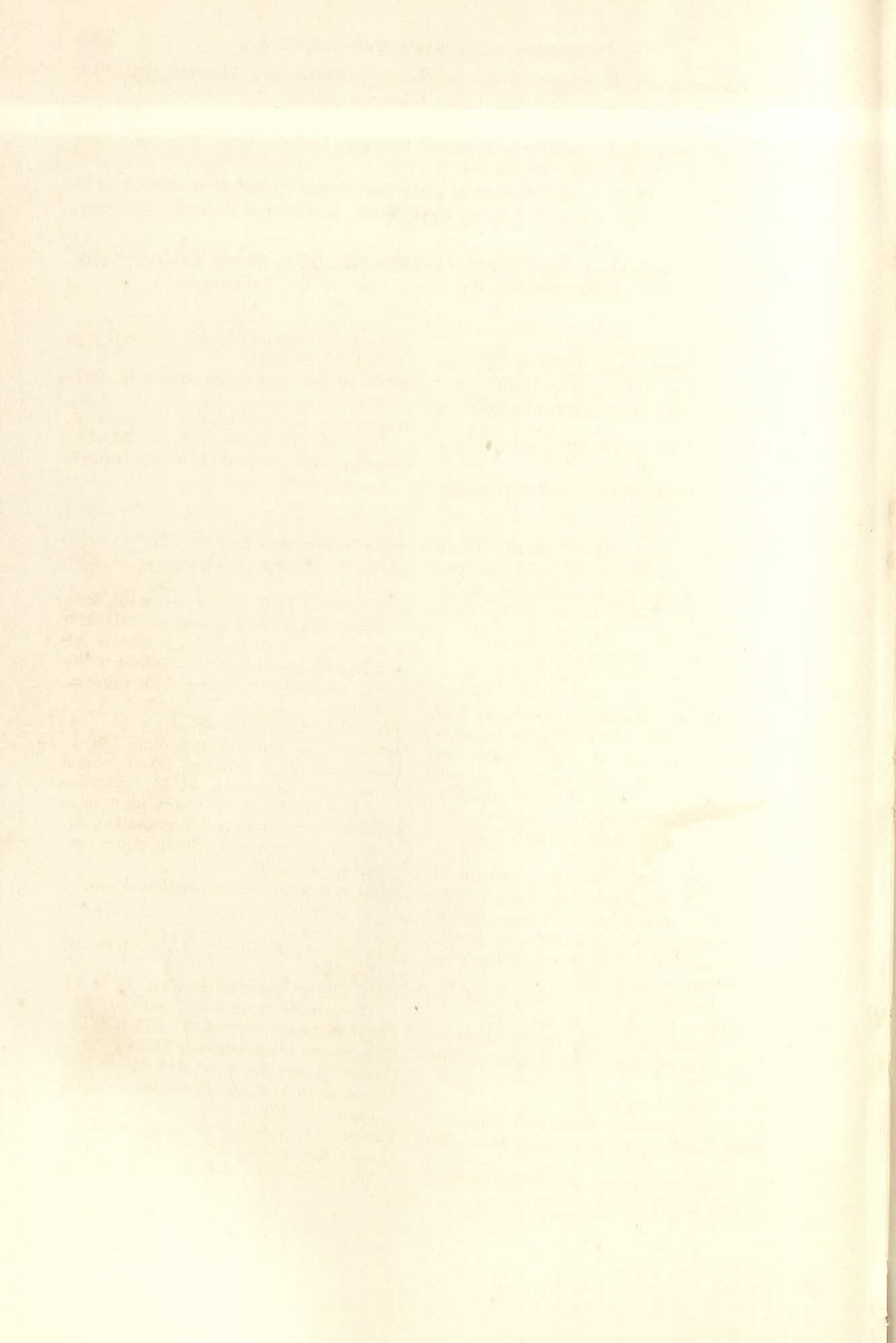
The psychophysical methods have aided in the comparison and grading of perfumes, liquors, tobaccos; and have even been applied to studies of public opinion. The method of average error, for example, has proved to be useful in studies of the accuracy with which blindfolded pilots were able to reproduce pressures ranging from one to forty pounds on a stick-type control. Psychometric functions like those of Figure 64 have been determined for the normal and hard of hearing. These curves show the relationship between intensity of speech and audibility, and are useful in vocations in which normal hearing is important.

In spite of the trend away from Fechner's psychophysics, as one reads through the experimental work in psychology since Fechner's time (and since James's, too) he cannot help being impressed with the influence exerted by psychophysics and the psychophysical methods. Fechner's work was important because it showed conclusively that problems in psychology can be subjected to quantitative methods and are amenable to exact mathematical treatment. The far-reaching result of this point of view is seen today in the development of mental, educational, and achievement tests and scales, and in the increasing use of statistical method in the study of human variability. Perhaps it is not too much to say that, beginning with Fechner, psychology ceased to be a branch of philosophy and began to emerge as an experimental science.

BIBLIOGRAPHY

1. Boring, E. G. A history of experimental psychology, 2nd ed. New York: Appleton-Century-Crofts, Inc., 1950, chapter 14.
2. Fitts, P. M. (Ed.) Psychological research in equipment design. Army Air Forces Aviation Psychol. Program, Res. Reports. U.S. Govt. Printing Office, 1947, 19, chapter 12.
3. Fletcher, H. Speech and hearing. New York: D. Van Nostrand Company, Inc., 1929, 68-70.
4. Fullerton, G. S., & Cattell, J. McK. On the perception of small differences. University of Pennsylvania Philosophical Series, 1892, No. 2.
5. Hecht, S. The nature of the photoreceptor process. In Murchison (Ed.), A. handbook of general experimental psychology. Worcester, Massachusetts: Clark University Press, 1934, chapter 14.
6. James, W. Principles of psychology. New York: Henry Holt and Company, 1890, I, p. 534.
7. Kellogg, W. N. An experimental evaluation of equality judgments in psychophysics. Arch. Psychol., New York, 1930, 17, No. 112.
8. Stevens, S. S., & Davis, H. Hearing: its psychology and physiology. New York: John Wiley & Sons, Inc., 1938.
9. Thorndike, E. L. Handwriting. Teach. Coll. Rec., 1910, 11, 1-81.
10. Thurstone, L. L. A method of scaling psychological and educational tests. J. educ. Psychol., 1925, 16, 433-451.

11. Titchener, E. B. Experimental psychology. Instructor's Manual, Quantitative. New York: The Macmillan Company, 1923.
12. Woodworth, R. S. Professor Cattell's psychological contributions. In The psychological researches of James McKeen Cattell. Arch. Psychol., New York, 1914, 4, No. 30, 60-74.
13. Woodworth, R. S. Psychological data pertaining to errors of observation. International Critical Tables. New York: McGraw-Hill Book Company, Inc., 1926, I, 92-95.
14. Woodworth, R. S. Experimental psychology. New York: Henry Holt and Company, 1938, chapter 18.



INDEX

NOTE. Proper names are not included unless indispensable in tracing a fact or theory. Instead of Pavlov, for example, see Conditioned reflex for detailed reference.

- Adrenin, in emotions, 152
- Age, and intelligence, 253-254; reaction time loss in, 321; *see also* Mental age
- Allport, F. H., list of prepotent reflexes, 201-202
- Allport and Vernon, studies in expressive movements, 191-192
- Ambiguous figures, in Gestalt psychology, 66-70
- Anger. *See* Rage
- Animal learning, and CR, 6-10; experiments in, 14-16, 41-56, 70-74; interpretations of, 44-46; reasoning in, 46-50; insight in, 50-52; imitation in, 52-53; principles of, 53-54
- Aphasia, brain injury in, 25-27
- Areas of brain, in learning, 22-24
- Army Alpha, description of, 244-245; distribution of soldiers' scores on, 250; limitations of, 250-251; scores of occupational groups on, 256; Negro and white scores on, 258; foreign-born men compared on, 260; effect of schooling on, 261-262; college success and, 262
- Army Beta, description of, 245, 248
- Army General Classification Test (AGCT), description of, 245-246; content of, 248-249; distribution of scores on, 252; comparison with Army Alpha, 252; occupational level and, 254-256; predictive value of, 258; Negro-White comparisons on, 259-260
- Army Individual Test of General Mental Ability, description of, 231-232
- Aspiration level, studies of, 202-203
- Assessment of Men, OSS Project in, 213-214
- Association, laws of, 2-3
- Associative reaction time, 323-324
- Attitude, effect of, in reaction time, 321-323
- Attitudes, studies of, 187-188; scales for the measurement of, 187; in social psychology, 209-211
- Auditory perception, apparatus in, 292-294
- Autonomic nervous system, divisions of, 151-152; rôle in emotions of, 152-154
- Average error, method of, in psychophysics, 345-348
- Behaviorism, definition of, 127
- Binet, A., concept of intelligence of, 220-221
- Binet-Simon Scale, construction of, 220-223; 1911 scale reproduced, 221-222; revisions of, 223-225; criticisms of, 232-235; limitations of, 233-234; training necessary for giving, 235; value of, 240-241
- Binocular vision, 302-307
- Boring, E. G., quoted on Galton, 289
- Brain, rôle of, in learning, 20-27; operation upon, in animals, 21-32; restitution of function in, 24-27; injury to, in man, 25-26; in brightness discrimination, 27-29; in motor habits, 29-32; equipotentiality of, 32-35
- Brain waves. *See* Electroencephalography
- Cannon, W. B., theory of emotions of, 151-154
- Cats, rôle of brain in learning of, 21-24; Thorndike's experiments with, 43-46
- Child behavior, studies of, 127-131
- Chimpanzees, learning by, 71-74
- Choice, in reaction time, 315-316
- Chronoscope, use in measuring reaction time, 316-318
- Closure, in Gestalt psychology, 68-70
- Conditioned reflex, definition of, 2; Pavlov's method in, 3-4; how established,

- Conditioned reflex (*continued*)
 4-5; time factor in, 5-7; differentiation and generalization in, 7-8; inhibition and extinction in, 8-10; experimental neurosis and, 10-12; and sleep, 12-14; in animals and children, 14-16; in adults, 16; as a principle of learning, 17
- Conditioned response (CR), definition of, 2; Pavlovian and instrumental, 14; illustrations of, in learning, 14-16; contrasted with associative shifting, 54-56; illustrations of, in emotion, 139-143
- Configuration, in Gestalt psychology, 64
- Conforming behavior, studies of, 205
- Constant stimuli, method of, in psychophysics, 342-345; simpler form of, 346
- Correlation, in the study of individual differences, 269-273
- Co-twin control, experimental method in, 132-133
- Delayed response, 49
- Delinquency, value of mental tests in, 233-234
- Development, motor and sensory, in young children, 128-131; mental and social, 131-133; tests of, 135-138; emotional, 139-140
- Diagnostic value, of intelligence tests, 231-232, 240-241
- Differential psychology, definition of, 265-266
- Differentiation, in conditioned reflex, 7-8; loss of, 10-12
- Discrimination, in conditioned reflex, 7-10; elimination by subtraction in, 315; in reaction experiments, 315, 324, 326
- Discrimination time method, in testing reaction time, 324-326; use by Air Forces in selection, 328
- Distribution versus concentration in learning, 118, 120-121; of intelligence, 227; of human traits, 266-269
- Economy, in memorizing, 120-124
- Effect, law of, 58-59; present status of, 59-61; criticisms of, 60; evaluation of, 61
- Electroencephalography, the study of brain waves, 35-37; alpha, beta, and delta waves, 36; factors in production of, 36-37
- Emotional development, in children, 139-140; experiments on, 140-143; methods of studying, 143-146
- Emotions, theories of, 148-150; problems and methods in studying, 150-151; physical basis of, 151-154; physiological patterns in, 154-157; free association in, 162-165; and deception, 166-168; present status of study of, 168-169
- Environment, influence of, upon intelligence, 232-235, 273-279
- Equipotentiality and mass action, in brain, 32-35
- Exercise, law of, 56-58; sub-laws under, 57-58
- Experimental method, in social psychology, 207-209
- Expression, method of, in emotions, 150
- Extinction, experimental, 8-10; and forgetting, 10
- Eye movements, nature of, in perception, 307-309; in reading, 309-310
- Facial expression, in emotion, 157-159; as judged by children, 159
- Factorial studies of personality organization, 193-195
- Faculty psychology, defined, 83-84; evidence against, 85-87
- Fear, as primary emotion, 139; acquisition of, 140-143
- Feeble-mindedness, occurrence in population, 228; nature of, 229; inheritance of, 237-238
- Figure and ground, in Gestalt psychology, 66-69
- Filial regression, Galton's law of, 271
- Forgetting, and extinction, 10; curve of, 111; and passing of time, 115-116
- Formal discipline, and faculty psychology, 83-87; illustrations of belief in, 84-85
- Foreign-born, intelligence of, 260
- Foster children, studies of, 275-277
- Free association, 162-165; detective use of, 166-168
- Freud, S., influence upon study of personality, 172-174; view of child growth, 173
- Frustration-aggression hypothesis, studies of, 205-207
- Galton, Francis, pioneer work of, on individual differences, 264-275
- Genius, in childhood, 235-237

- Gestalt psychology, meaning of, 63-64; experiments in perception, 64-70; experiments in learning, 70-76; Lewin's work in, 75-76; studies of, in America, 76-78; present status of, 78-81
- Gifted children, Terman's study of, 229, 235
- Glandular theories of personality, 180-181
- Graphology, in study of personality, 181
- Group mind, definition of, 204
- Group tests of intelligence, construction of, 246-249; effect of age upon, 253-254; relation to occupational level, 254-257; value in education, 262
- Growth, changes over first five years, 135-138; curve of, for general intelligence, 226
- Habits, as CR's, 14-16; methods of breaking undesirable, 143-146
- Halo effect, in ratings, 183
- Handedness, studies of, in children, 133-134
- Hearing, theories of, 296-297
- Helmholtz, H. von, theory of hearing, 296
- Heredity and environment, 232-239, 272-279
- Honesty, studies of, in children, 189-191
- Ideas, in learning, 46-52
- Imagery, types of, 266-267
- Imitation, in animals, 52-53
- Impression, method of, in study of emotions, 150
- Impressionistic methods, vogue of, in study of personality, 181-182
- Impulse, in emotion, 149, 169
- Individual differences, early tests of, 264-266; meaning of, 266-269; nature and nurture in, 272-273, 275-283
- Individual intelligence tests, criticisms of, 232-235; evaluation of, 235
- Infants, motor and sensory development in, 128-131
- Inference, in learning, 46-52
- Inheritance, statistical studies of, 273-275; experimental studies of, 275-281
- Inhibition, in conditioned reflexes, 6-10; and extinction, 8-10; in sleep, 12-14; retroactive and proactive, 105
- Insight, in animal learning, 50-52; in Gestalt psychology, 74; criticisms of, 79-80; in educational practice, 80-81
- Instincts, McDougall's list of, 200; in social psychology, 200-202; criticisms of, 201
- Intelligence, Binet's concept of, 220-221; growth of, 225-226; maturity of, 226; distribution of, 227-229; testing of, in adults, 229-232; effect of age upon, 231, 253-254; native basis of, 235-238; of the soldier, 249-252; and occupational level, 254-257; racial differences in, 257-260; of national groups, 260-261; and schooling, 261-262; sex differences in, 282-283
- Intelligence quotient (IQ), 224; distribution in population, 227; constancy of, 239-240; causes of variations in, 239-240; and school success, 240-241
- Introversion-extroversion, studies of, 193-195
- James-Lange theory of emotions, 149; criticisms of, 152-153; substitute for, 153
- J-Curve, definition of, 205-206
- Jung, Carl, experiments of, 162-164, 166-167
- Kallikak family, history of, 237-239
- Kent and Rosanoff, free-association test, 162-165
- Krasnogorski, N., studies of conditioned reflex, 15-16
- Kuhlmann, F., tests for young children, 135; revision of Binet scale by, 223
- Latent time, definition of, 6
- Latin, belief in disciplinary value of, 91-93
- Laws of learning, Thorndike's, 56-61
- Learning, and CR, 6-10; in animals, 14-16, 41-56, 70-74; in man, 16-17, 55-59, 76, 78; Thorndike's principles of, 53-54, 56-59; in Gestalt psychology, 70-78; and retention in memory, 106-109; repeated, effect of, 106-109; overlearning and review in, 106-109; distributed versus concentrated, 113-118; whole versus part, 120-122; active versus passive, 122-123
- Length of material, and memory, 104-106
- Lewin, K., experiments in motivation, 74-76

- Localization, in brain, 22-23
 Locke, John, quoted on association, 2-3
 Love, as primary emotion, 139
- Masculinity and femininity, studies of, 192-193
 Mateer, F., work of, upon conditioned response, 16
 Mathematics, belief in disciplinary value of, 93-95
 Maturity of intelligence, 226
 Maze, use of in learning, 41-43; in transfer experiments, 90-91
 McDougall, William, contributions to social psychology, 199-201; list of instincts, 200
 Memorizing, problems in, 103; Ebbinghaus's standard of, 104; amount and effort in, 104-106; effect of time upon, 110-112; and forgetting, 111, 117, 120-124; influence of method upon, 112-115; serial connections in, 118-120; backward association in, 120; economy and efficiency in, 120-124
 Memory, phenomena of, 102-103; threshold of, 116; training in, 120-124
 Mental age (MA), 221; of soldiers as measured by Army Alpha, 250-251
 Mental growth, in young children, 135-138; curve of, 226
 Mental organization, modern view of, 85-87
 Merrill-Palmer tests, for young children, 135-136
 Minimal change, method of, 341-342
 Minnesota preschool tests, 136
 Müller-Lyer illusion, 306
 Murray, H. A., studies in personality by, 195-196
- National groups, intelligence of, 260, 287-289
 Negroes, attitudes toward Negro soldiers, 214-216; intelligence of, 257-260; Northern and Southern, 258; comparison, white and Negro children, 285-286; studies of temperamental traits in, 287
 Nerve conduction, speed of, 313-314
 Neurosis, experimental, and conditioned reflex, 10-12
 Nonsense syllables, value of, in memory studies, 101-102
 Normal curve, of IQ's, 227
- Observation, as method in social psychology, 207-209
 Occupation, and intelligence, 254-257
 Opinion-polling, methods in, 207-209; importance of questions in, 208; studies of, 214-216
 Organization, principles of, in Gestalt psychology, 69-70
 Overlapping, of races in intelligence, 257-260, 283-287; of sexes, 279-283; meaning of, 280
 Overlearning, meaning of, 106-109; 122-123
- Palmistry, as pseudo-science, 181
 Patterns, in learning, 70-72; in behavior, 139-140
 Pearson, K., statistical studies of, 271, 274-275
 Perception, organization in, 64-70; importance of, 292; experience and learning in, 297-302; physiological factors in, 302-304; psychological factors in, 304-307
 Performance tests, illustrations of, 137
 Personality, definition of, 171-172; factors in, 174-175; impressionistic methods of studying, 175-181; experimental methods of studying, 183-189; tests in, 188-189; experimental investigations of, 189-197
 Phi-phenomenon, meaning of, 64-66
 Phrenology, and faculty psychology, 86-87; in the study of personality, 176-177
 Physiognomy, in the study of personality, 176-180
 Pintner-Paterson Scale of Performance Tests, 137-138
 Poetry, forgetting in, 112
 Pregnancy, law of, 68-70
 Prestige suggestion, studies of, 202
 Primary emotions, 139
 Primitive folk, abilities of, 284-285
 Principles of learning, illustrations of, 41, 43, 49, 54-56, 57-58; Thorndike's, 53-54, 56-59
 Profile, representation of tests by, 231-232
 Propaganda, studies in, 211-212
 Pseudoscope, in studies of perception, 301-302
 Psychoanalysis, experimental studies of, 173-174
 Psychogalvanic reflex, description of, 154; in studies of emotion, 159-162

- Psychophysics, definition of, 331-332; methods in, 340-348; present status of, 349-350
- Psychosurgery, effects of, in mental patients, 37-38
- Questionnaire, in personality study, 186-188
- Quotient, intelligence, meaning of, 224
- Race differences, Negro-white, in America, 257-260; in mental ability, 283-289
- Rage, as primary emotion, 139
- Rating scale, in personality study, 183-185
- Rats, rôle of brain in learning of, 27-32
- Reaction time, in psychology, 312-313; history of, 313-314; early studies of, 314-316; simple, discriminative, and choice, 314-316; apparatus used in measurement of, 316-318; associative, 318; factors influencing, 318-321; effect of attitude upon, 321-323; applications of, 326-329; use of, in World War II for selection, 328
- Reading, eye movements in, 309-310; effect of reaction time studies upon, 324
- Recitation, value of, in memorizing, 116-118
- Reflexes, definition of, 2; Babinsky, 130; prepotent reflexes, list of, 201-202
- Resemblances, family, in mental traits, 272-273; criticisms of studies of, 273-275
- Restitution of function, in brain injury, 24-27; organization in, 32-35
- Retention, curve of, 111; for nonsense syllables, 111-112; in various school subjects, 115-116; effect of review upon, 116-118
- Rôle or status, in social behavior, 204-205
- Rorschach test, in the study of personality, 192
- Salivary reflex, in Pavlov's method, 3-5
- Satisfaction, in learning, 58-59
- Savings method, in memory, 110
- School learning, transfer in, 91-95
- Seashore tests of musical talent, description of, 293-294
- Sentiments, definition of, 200
- Serial association, in learning, 118-120
- Sex differences, in abilities and traits, 279-283; summary of, 283
- Sheldon, W. H., type theory of, 178-180
- Sleep, and inhibition, Pavlov's theory of, 12-14
- Social behavior, factors in, 202-207; systematic studies of, 209-217
- Social facilitation, studies of, 203-204
- Social psychology, definition of, 199; methods in, 207-209
- Sociometric technique, in social psychology, 210-211
- Sounds, characteristics of, 294; spatial localization of, 294-296
- Space perception, theories of, 297-298, 310
- Stanford Revision (Stanford-Binet), construction of, 223-224; 1937 scale, 224-225; administration of, 235
- Stereoscope, Wheatstone, 298; Brewster, 298-301
- Stereotypes, studies of, 209-211
- Stern, W., definition of IQ, 224
- Stratton, G. M., experiment of, upon eye movements, 308-310
- Suggestion, as incentive in brain injury, 25-27
- Superior children, Terman's study of, 229, 235
- Sympathetic branch of the nervous system, rôle in emotion of, 151-154
- Tests of development in young children, 135-138
- Thalamus, rôle of, in emotion, 153-154
- Thinking, and problem solving, in animals, 46-53; in man, 55-56
- Thorndike, E. L., laws of, 53-54, 56-60
- Topological studies of personality, 196-197
- Transfer of training, James's experiment on, 87-88; Woodworth's and Thorndike's experiments on, 88-90; in maze learning, 90-91; in school learning, 91-95; theories of, 95-96; summary of, 96-97; techniques in the study of, 97-99
- Transpositional learning, meaning of, 70-72
- Trial-and-error, learning by, 41-46; and insight, 50-53
- Twins, in studies of nature and nurture, 277-279
- Types, psychological theories of, 176-181; body, 178-180; personality, 178-180

- Use, law of, 56-58
- Validity, of intelligence tests, 245
- Variability in mental traits, male and female, 282-283
- Visual perception, Gestalt studies of, 64-70; importance of, 292; of distance and depth, 297-307; retinal image in, 298; in three dimensions, 299; physiological factors in, 302-304; psychological factors in, 304-307
- Vocational selection, 230-232; interests and, 187-188; reaction time in studies of, 328
- Wallin Peg Boards, 137
- Weber-Fechner law, description of, 334-337; limitations of, 337-339; interpretations of, 339-340
- Weber's law, 332-334
- Wechsler-Bellevue Scale, in measurement of adult intelligence, 230-231; IQ in, 230; diagnostic value of, 231-232
- Wechsler Intelligence Scale for Children, description of, 230
- Wever, E. G., and Bray, C. W., experiment of, in audition, 297
- Worker productivity, social factors in, 216-217
- Yerkes, R. M., revision of Binet scale by, 223
- Yerkes discrimination box, in animal learning, 47-49

THE CENTURY PSYCHOLOGY SERIES

Richard M. Elliott, *Editor*

Kenneth MacCorquodale, *Assistant Editor*

- Social Psychology*, by Charles Bird
Learning More by Effective Study, by Charles and Dorothy Bird
Psychological Counseling, by Edward S. Bordin
A History of Experimental Psychology, 2d ed., by Edwin G. Boring
Sensation and Perception in the History of Experimental Psychology, by
Edwin G. Boring
Readings in Modern Methods of Counseling, edited by Arthur H. Bray-
field
A Casebook of Counseling, by Robert Callis, Paul C. Polmantier, and
Edward C. Roeber
Beauty and Human Nature, by Albert R. Chandler
Readings in the History of Psychology, edited by Wayne Dennis
Techniques of Attitude Scale Construction, by Allen L. Edwards
Modern Learning Theory, by William K. Estes, Sigmund Koch, Kenneth
MacCorquodale, Paul E. Meehl, Conrad G. Mueller, Jr., William N.
Schoenfeld, and William S. Verplanck
Schedules of Reinforcement, by C. B. Ferster and B. F. Skinner
Social Relations and Morale in Small Groups, by Eric F. Gardner and
George G. Thompson
Great Experiments in Psychology, 3d ed., by Henry E. Garrett
Exceptional Children, by Florence L. Goodenough
Developmental Psychology, 3d ed., by Florence L. Goodenough and
Leona E. Tyler
Physiological Psychology, by Starke R. Hathaway
Seven Psychologies, by Edna Heidbreder
Theories of Learning, 2d ed., by Ernest R. Hilgard
Conditioning and Learning, by Ernest R. Hilgard and Donald G. Marquis
Hypnosis and Suggestibility, by Clark L. Hull
Principles of Behavior, by Clark L. Hull
Development in Adolescence, by Harold E. Jones
The Definition of Psychology, by Fred S. Keller
Principles of Psychology, by Fred S. Keller and William N. Schoenfeld
Psychological Studies of Human Development, by Raymond G. Kuhlen
and George G. Thompson
The Cultural Background of Personality, by Ralph Linton

- Vocational Counseling with the Physically Handicapped*, by Lloyd H. Lofquist
- Studies in Motivation*, edited by David C. McClelland
- The Achievement Motive*, by David C. McClelland, John W. Atkinson, Russell A. Clark, and Edgar L. Lowell
- Current Studies in Psychology*, by F. Joseph McGuigan and Allen D. Calvin
- Principles of Applied Psychology*, by A. T. Poffenberger
- The Behavior of Organisms: An Experimental Analysis*, by B. F. Skinner
- Cumulative Record*, by B. F. Skinner
- Verbal Behavior*, by B. F. Skinner
- Diagnosing Personality and Conduct*, by Percival M. Symonds
- Dynamic Psychology*, by Percival M. Symonds
- The Dynamics of Human Adjustment*, by Percival M. Symonds
- The Ego and the Self*, by Percival M. Symonds
- The Psychology of Parent-Child Relationships*, by Percival M. Symonds
- Educational Psychology*, by George G. Thompson, Eric F. Gardner, and Francis J. DiVesta. Also accompanying Workbook by the same authors.
- Selected Writings from a Connectionist's Psychology*, by Edward L. Thorndike
- Introduction to Methods in Experimental Psychology*, 3d ed., by Miles A. Tinker and Wallace A. Russell
- The Psychology of Human Differences*, 2d ed., by Leona E. Tyler
- The Work of the Counselor*, by Leona E. Tyler
- Experimental Psychology: An Introduction*, by Benton J. Underwood
- Psychological Research*, by Benton J. Underwood
- Elementary Statistics*, by Benton J. Underwood, Carl P. Duncan, Janet A. Taylor, and John W. Cotton. Also accompanying Workbook by the same authors.
- Persons and Personality*, by Sister Annette Walters and Sister Kevin O'Hara

Vol. 41, No. 1, January 1, 1929

Published by the American Medical Association, 535 North Dearborn Street, Chicago, Ill.

Subscription price, \$5.00 per annum in advance. Single copies, 15 cents.

Entered as second-class matter, June 26, 1911. Postpaid at special rate of \$3.75 per annum provided for by Act of October 3, 1917.

Acceptance for mailing at special rate of postage provided for in Act of October 3, 1917, authorized on July 1, 1928.

Postmaster: This journal is published weekly except on Sundays, and is mailed at special rate of postage provided for in Act of October 3, 1917.

Copyright, 1929, by American Medical Association. All rights reserved.

Printed at the American Medical Association Press, 535 North Dearborn Street, Chicago, Ill.

Second-class postage paid at Chicago, Ill., and at additional mailing offices.

Postpaid at special rate of postage provided for in Act of October 3, 1917.

Entered as second-class matter, June 26, 1911. Postpaid at special rate of postage provided for in Act of October 3, 1917.

Acceptance for mailing at special rate of postage provided for in Act of October 3, 1917, authorized on July 1, 1928.

Postmaster: This journal is published weekly except on Sundays, and is mailed at special rate of postage provided for in Act of October 3, 1917.

Copyright, 1929, by American Medical Association. All rights reserved.

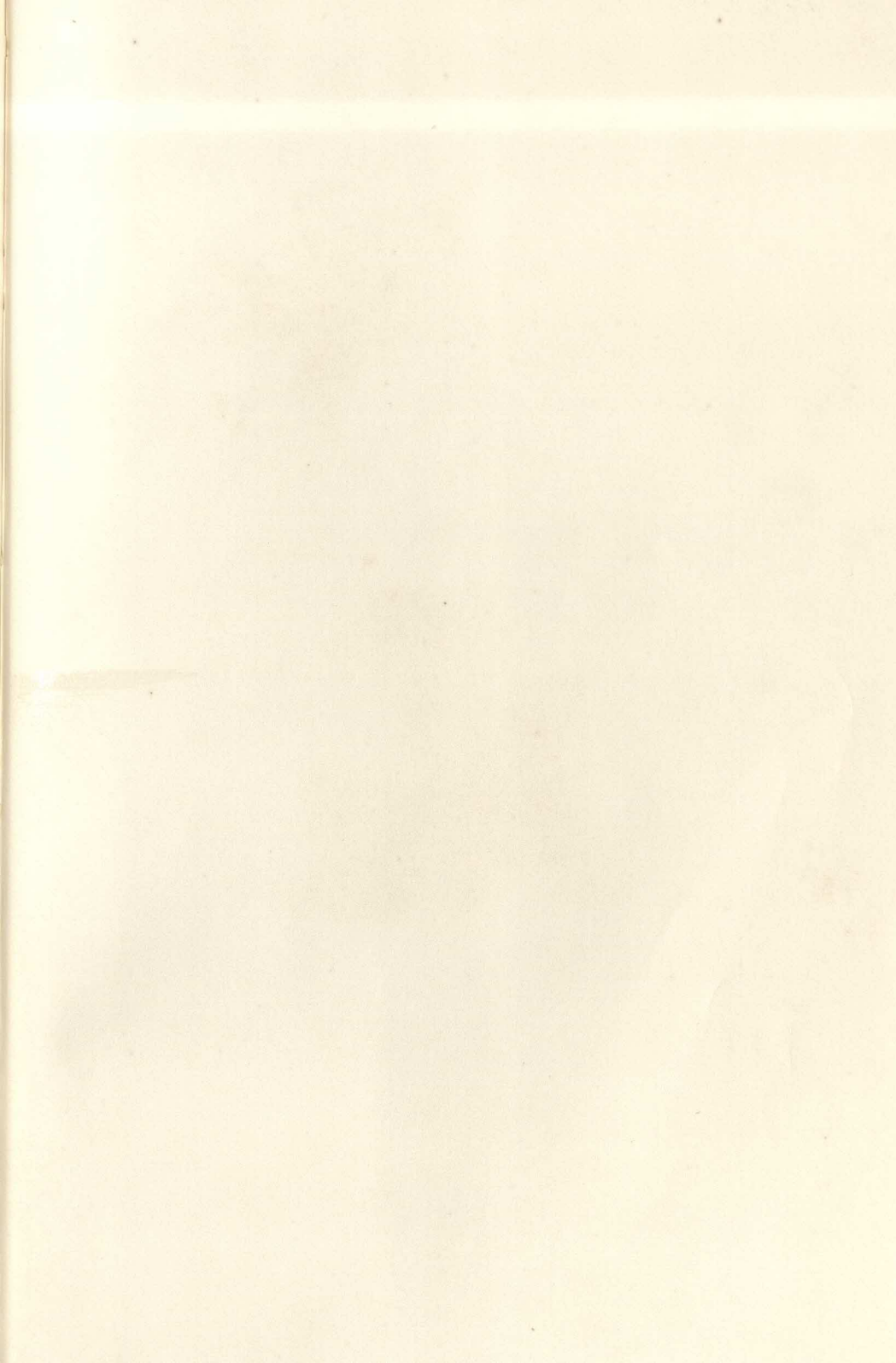
Printed at the American Medical Association Press, 535 North Dearborn Street, Chicago, Ill.

Second-class postage paid at Chicago, Ill., and at additional mailing offices.

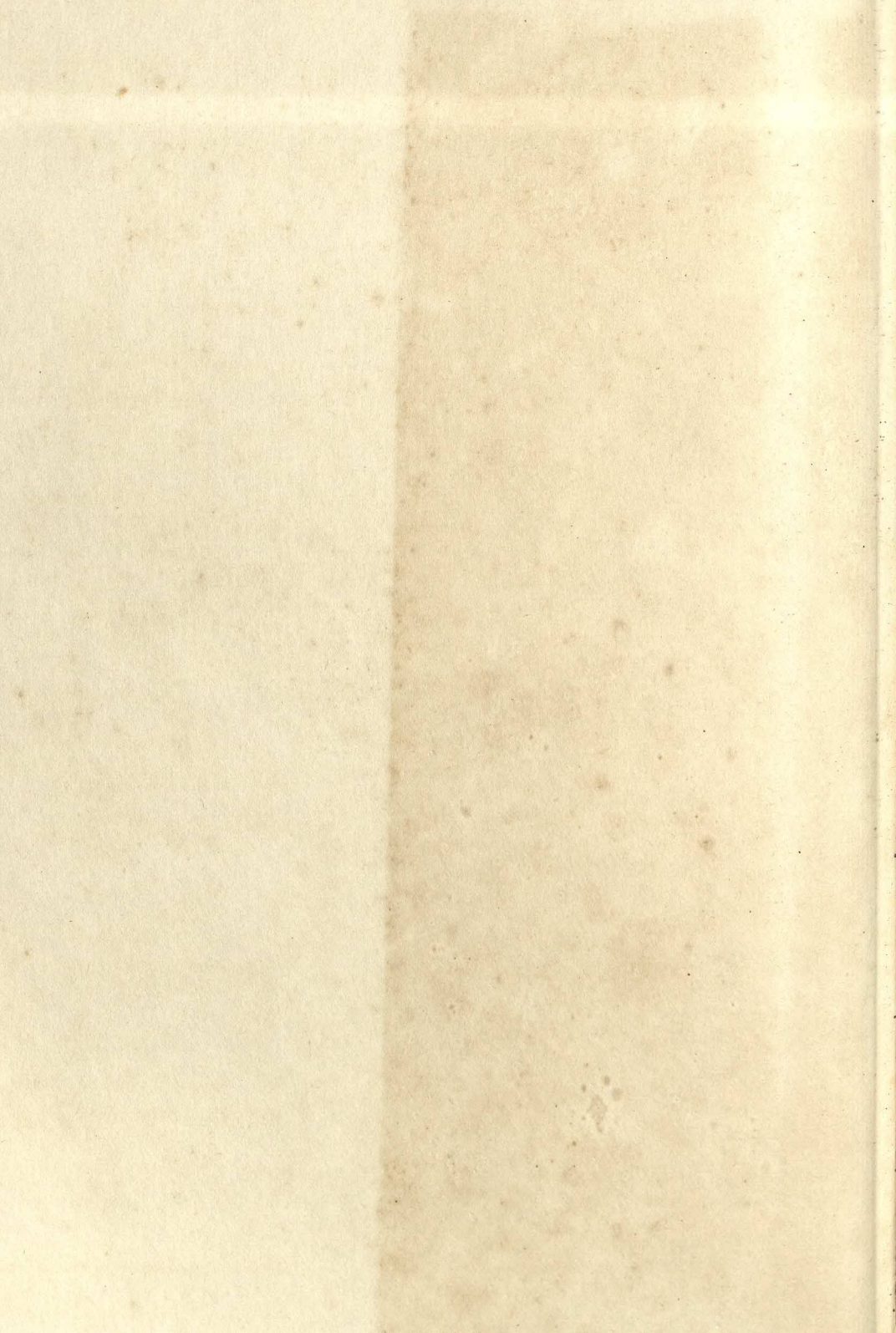
Postpaid at special rate of postage provided for in Act of October 3, 1917.

Entered as second-class matter, June 26, 1911. Postpaid at special rate of postage provided for in Act of October 3, 1917.

Acceptance for mailing at special rate of postage provided for in Act of October 3, 1917, authorized on July 1, 1928.







VISION